



USAISEC

US Army Information Systems Engineering Command
Fort Huachuca, AZ 85613-5300



U.S. ARMY INSTITUTE FOR RESEARCH
IN MANAGEMENT INFORMATION,
COMMUNICATIONS, AND COMPUTER SCIENCES

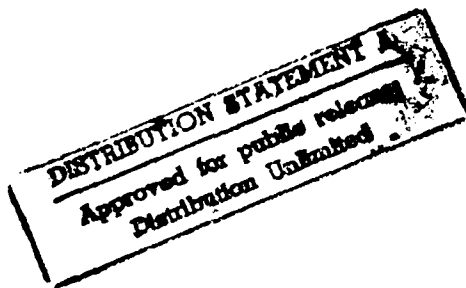
AD-A267 848



APPLICATION OF NEURAL NETWORKS FOR THE EXTRACTION
& CHARACTERIZATION OF KNOWLEDGE CONTAINED IN
DATABASES

ASQB-GM-91-013
JUNE 1991

DTIC
ELECTE
S AUG 12 1993 D
B



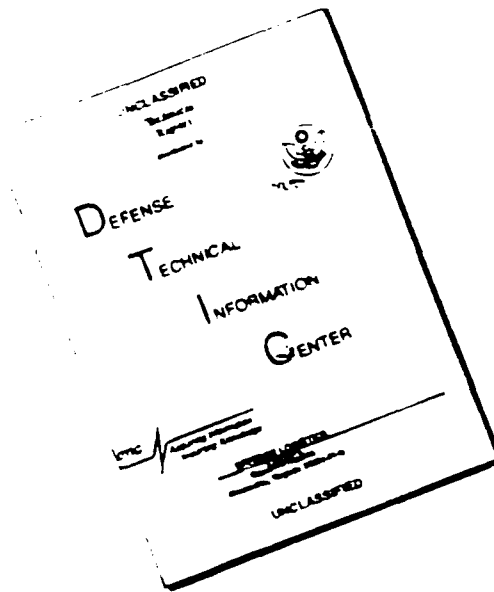
93-18405

AIRMICS
115 O'Keefe Building
Georgia Institute of Technology
Atlanta, GA 30332-0800

93



DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188
Exp. Date: Jun 30, 1986

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b. RESTRICTIVE MARKINGS NONE	
2a. SECURITY CLASSIFICATION AUTHORITY N/A			3. DISTRIBUTION/AVAILABILITY OF REPORT N/A	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A			5. MONITORING ORGANIZATION REPORT NUMBER(S) N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) ASQB-CM-91-013			7a. NAME OF MONITORING ORGANIZATION N/A	
6a. NAME OF PERFORMING ORGANIZATION AIRMICS		6b. OFFICE SYMBOL (If applicable) ASQB-CM		7b. ADDRESS (City, State, and ZIP Code) N/A
6c. ADDRESS (City, State, and Zip Code) 115 O'Keefe Bldg. Georgia Institute of Technology Atlanta, Ga 30332-0800			9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	
6d. NAME OF FUNDING/SPONSORING ORGANIZATION AIRMICS		6e. OFFICE SYMBOL (If applicable) ASQB-CM		10. SOURCE OF FUNDING NUMBERS
6f. ADDRESS (City, State, and ZIP Code) 115 O'Keefe Bldg. Georgia Institute of Technology Atlanta, GA 30332-0800			PROGRAM ELEMENT NO. 62783A	PROJECT NO. DY10
			TASK NO. 05	WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) Application of Neural Networks for the Extraction & Characterization of Knowledge Contained in Databases				
12. PERSONAL AUTHOR(S) William R. Caid				
13a. TYPE OF REPORT		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Year, Month, Day) 1991, March, 26
15. PAGE COUNT 83				
16. SUPPLEMENTARY NOTATION				
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUBGROUP	Data Base Mining, Neural Nets, Relationship Discovery, Data Modeling, Mapping Networks, Self Organization	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) The ability to automatically discover relationships contained within data, quantify their strength, and present them graphically to the user for visualization is defined as "Relationship Discovery". This capability was the major research effort during Phase I of this SBIR Project. The detection of relationships is a necessary precursor to the modeling step where the detected relationships are modeled using powerful neural network paradigms capable of capturing the nonlinear relationships in the data. As a result of this Phase I SBIR Project, a relationship discovery capability has been developed to automatically determine the existence of a relationship in each sub-space and to determine the strength of the relationship. This capability is then used to produce a comprehensive listing of variables that are related to other variables and the strength of the relationship. As a result of the development and testing of these concepts on data sets, several necessary enhancements to the overall Data Base Mining system concept were identified. These are Missing Value Prediction, Bad Data Detection, and Data Redundancy[1]				
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Michael E. Mizell			22b. TELEPHONE (Include Area Code) 404/894-3107	22c. OFFICE SYMBOL ASQB-CM

[1]

Removal.

The power of the Relationship Discovery approach is that it makes no assumptions about the statistics of the data contained within the database. Additionally, this approach makes no assumptions about how many relationship regions exist in each sub-space. These powerful characteristics allow the relationship discovery concept to be applied to a wide range of problems.

This research was performed for the Army Institute for Research in Management Information, Communications and Computer Sciences (AIRMICS), the RDTE organization of the U.S. Army Information Systems Engineering Command (USAISEC). This report is not to be construed as an official Army position, unless so designated by other authorized documents. Material included herein is approved for public release, distribution unlimited. Not protected by copyright laws.

THIS REPORT HAS BEEN REVIEWED AND APPROVED

s/ James D. Gantt
 James D. Gantt
 Division Chief
 MISD

s/ John R. Mitchell
 John R. Mitchell
 Director
 AIRMICS

DTIC QUALITY INSPECTED 3

- 1 -

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

FINAL REPORT

Contract No. DAKF11-90-C-0036

March 26, 1991

Prepared By:

**HNC, Inc.
5501 Oberlin Drive
San Diego, CA 92121**

Table of Contents

Executive Summary.....	1
Introduction	1
Objectives of Phase.....	1
Results of Phase I	2
Necessary Enhancements Identified in Phase I	2
Summary of Results Obtained in Phase I.....	3
 Description of End-Use Problem.....	 4
 Introduction to Relationship Discovery.....	 7
 Decreasing the Size of the Problem	 8
 Approaches to Relationship Discovery.....	 11
Relationship Discovery Via the Chi-squared Test	11
Relationship Discovery Via Probing of K-space.....	13
 Relationship Aggregation	 15
 Integrated Relationship Discovery Tool.....	 17
 Testing.....	 18
Tests on Chi-squared Relationship Discovery	18
Tests on Probing-of-K-space Relationship Discovery	19
Tests using Civilian Customer Data	19
 Extensions.....	 20
Necessary Enhancements Identified in Phase I	20
Missing Value Prediction	20
Bad Data Detection	26
Data Redundancy Removal	30
 Conclusions and Discussion.....	 35
Results of Phase I	35
Necessary Enhancements Identified in Phase I	36
Summary of Results Obtained in Phase I.....	36
 Technical References.....	 38
 Appendix 1 — Chi-squared Relationship Discovery on Test Data Set 1	
Appendix 2 — Chi-squared Relationship Discovery on Test Data Set 2	
Appendix 3 — Probing of K-space on Test Data Set 1	
Appendix 4 — Probing of K-space on Test Data Set 2	
Appendix 5 — Chi-squared Relationship Discovery on Civilian Customer Data	

Executive Summary

Introduction

The ability to automatically discover relationships contained within data, quantify their strength, and present them graphically to the user for visualization is defined as "Relationship Discovery". This capability was the major research effort during Phase I of this SBIR Project.

Relationship Discovery is the first step in HNC's proprietary Data Base Mining process. It determines "what is important" in the data and estimates the strength of the relationships between the variables. This detection of relationships is the necessary precursor to the modeling step where the detected relationships are modeled.

The approach uses a self organizing neural network technique to approximate the probability density function (PDF) of the data set. This approximation of the PDF is necessary on large data sets to reduce the size of the problem and make it computationally tractable. The two-dimensional projections of the PDF are then automatically examined for the existence of relationships and a relationship strength value is assigned to the projection. The relationships with the highest strength are presented either rapidly as a two-dimensional scatter plot or as a three-dimensional graph of the PDF, with the Z dimension being the amplitude of the PDF at each point.

This approach is similar to correlation analysis, but provides more accurate and useful results since it can accommodate nonlinear as well as linear relationships. Additionally, this approach can handle relationships that are discontinuous or exist only over a limited range of values. Unlike conventional statistical approaches based on correlation analysis, the proposed Relationship Discovery technique makes no assumptions about the nature of any functional or nonfunctional relationships contained within the data. Instead, the capability is driven by the underlying probability density of the data, thereby giving a true representation of any relationships.

Objectives of Phase I

The Phase I effort for this SBIR concentrated on the development of the automated relationship discovery and relationship aggregation techniques. The objectives for Phase I are listed below:

- Refine and expand the analysis techniques used for the determination of the existence of a relationship between specific variables.
- Automate the relationship determination techniques.
- Quantify the "strength" of relationships and develop a mechanism to ignore relationships that have less than an automatically determined cutoff strength.
- Develop an automated process to aggregate the relationships and build relationship hierarchies.
- Develop proof-of-concept code that implements the above techniques.
- Perform tests of the code on simulated and real data.

Appendix 1 — C++ Squared
Appendix 2 — C++ Squared
Appendix 3 — Prob
Appendix 4 — Prob
Appendix 5 — C++ Squared

Results of Phase I

All of the Phase I technical objectives were met and alternate approaches to the relationship discovery process were developed. Furthermore, a significant amount of refinement on the relationship discovery tool was performed. This resulted in a fully functional capability that is currently in use at HNC. The Relationship Discovery software tool along with an executable copy of the HNC Data Base Mining software will be delivered to the Army (AIRMICS) along with this Final Report.

As a result of Phase I, the relationship discovery capability has been enhanced to automatically determine the existence of a relationship in each sub-space and to determine the strength of the relationship. This capability is then used to produce a comprehensive listing of variables that are related to other variables and the strength of the relationship. Furthermore, the user may now easily see the most important relationships without having to exhaustively examine all relationships.

The relationship discovery tool can also produce an aggregated listing of minimum paths between selected database variables. The path length can be thought of as being the inverse of relationship strength. If a pair of variables have a strong relationship strength, then they have a "close" distance. Weakly related variables have a longer relationship distance. The Relationship Discovery tool can compute the path of minimum length between two variables. This information can be used to assist in the development of succinct data models as well as gaining additional insight into the variables and the data.

The power of the Relationship Discovery approach is that it makes no assumptions about the statistics of the data contained within the database. As such, this approach will perform equally well on data with gaussian or non-gaussian statistics. Additionally, this approach makes no assumptions about how many relationship regions exist in each sub-space. These powerful characteristics allow the relationship discovery concept to be applied to a wide range of problems. The output of Relationship Discovery is a three-dimensional plot of the PDF for each of the two-dimensional sub-space projections for visual analysis and a list of the variables that are "important".

Necessary Enhancements Identified in Phase I

During the attainment of these results and as a result of the testing of these concepts on data sets, several necessary enhancements to the overall Data Base Mining system concept were identified. These are:

- **Missing Value Prediction:** This capability will provide a way of systematically replacing missing values in the data set with maximum likelihood estimates of their true values.
- **Bad Data Detection:** This capability will help provide identification of potentially "bad" records in the data set.
- **Data Redundancy Removal:** This capability is necessary to optimize the model explanation capability and eliminate ambiguous results.

The first two of these, Missing Value Prediction and Bad Data Detection, together comprise Automatic Data Cleaning.

These enhancements identified during the Phase I effort are being proposed for implementation during Phase II. The technical approaches for these enhancements as well as their implications are detailed in the "Extensions" section of this report.

Summary of Results Obtained in Phase I

In summary, the following results were achieved during Phase I of this SBIR project:

- Implementation and demonstration of an automatic approach to relationship discovery via 2-D subspace projections.
- Development of a Chi-squared test to find relationships.
- Implementation of the capability for a scatter-plot display for quick visualization of relationships.
- Development of the capability to perform efficient Parzen windowing on the approximated PDF to generate 3-D surface plots of the PDF projections.
- Development of the capability to display relationships in rank space.
- Development of an approach to the model reduction problem via the shortest path method using the Chi-squared test and Cramer's coefficient to generate the Shortest Path Tree.
- Implementation of these capabilities in software and verification of software performance on real and artificial data.
- Enhancement of the general software capabilities of the Relationship Discovery tool:
 - Increase in utilization of neurocomputer processing capabilities to reduce processing time.
 - Enhancement of usability through implementation of a simple graphical user interface.
 - Integration of the software into a single coalesced environment.
- Conceptualization of approaches to three necessary enhancements to the Data Base Mining Concept.

Description of End-Use Problem

Analysis of the data contained in large databases can be a complex and time consuming process. HNC has developed KENN (Knowledge Extraction using Neural Networks), also known as Data Base Mining, as a system level approach to this analysis problem. This approach is a new neural network-based concept for extraction, characterization, and exploitation of knowledge contained within large databases. This approach utilizes HNC proprietary techniques, including several patented concepts.

KENN technology is in existence today and has been successfully employed on a prototype basis for the analysis of several diverse types of data. KENN is a powerful tool for the analysis of data, particularly when relationships between variables are non-linear. The neural network approach employed by KENN is closely related to advanced statistical techniques. These techniques are especially useful when the parametric form of the relationships in the data are unknown. It has been shown [1,2] that these network techniques provide a rich set of basis functions for the solution of a wide class of problems. It has further been proven that multi-layer feed-forward networks are universal approximators for arbitrary functions [3,4]. This background work has demonstrated that neural networks are a valid and promising vehicle for the modeling and analysis of database relationships.

In addition to the provable robustness of the neural network approach, there are additional benefits:

- The level of expertise required to use this approach is, in general, significantly lower than that required by conventional statistical approaches.
- The KENN approach will, at a minimum, produce results that are equivalent to conventional approaches.
- The KENN approach is capable of solving non-linear problems due to its use of neural network techniques.
- The cost of the analysis is lower than a conventional approach because KENN uses adaptive techniques which rapidly model the underlying relationships without specification of the parametric model or hypothesis testing. Neural networks "learn by example", and as such, no programming is required.

HNC has used the KENN approach for analysis of real data sets with outstanding results. This development effort has resulted in a set of capabilities that, with the proposed enhancements, are directly applicable to the EIS/ESS problem. A summary of the characteristics and attributes of HNC's approach are as follows:

- Provides a new way of analyzing, interpreting and understanding the contents of databases.
- Can be combined with conventional techniques to provide significant new capabilities.
- Finds and determines the structure of relationships. As such, KENN can discover unknown relationships.
- Characterizes relationships in both a qualitative and quantitative fashion, and can characterize large databases in a concise format with controllable precision.

- Allows visualization of relationships and provides outputs in both graphical and tabular formats.
- Explains relationships by identifying discriminators and provides a ranking of their importance.
- No prior knowledge of relationships within the data is required and domain expertise is not required for successful analysis.
- Requires no programming.
- HNC currently provides this capability in a PC and Sun environment. Other workstation platforms can also be utilized. System performance is enhanced by utilizing a neural network co-processor such as HNC's commercial products, the ANZA Plus and Balboa.

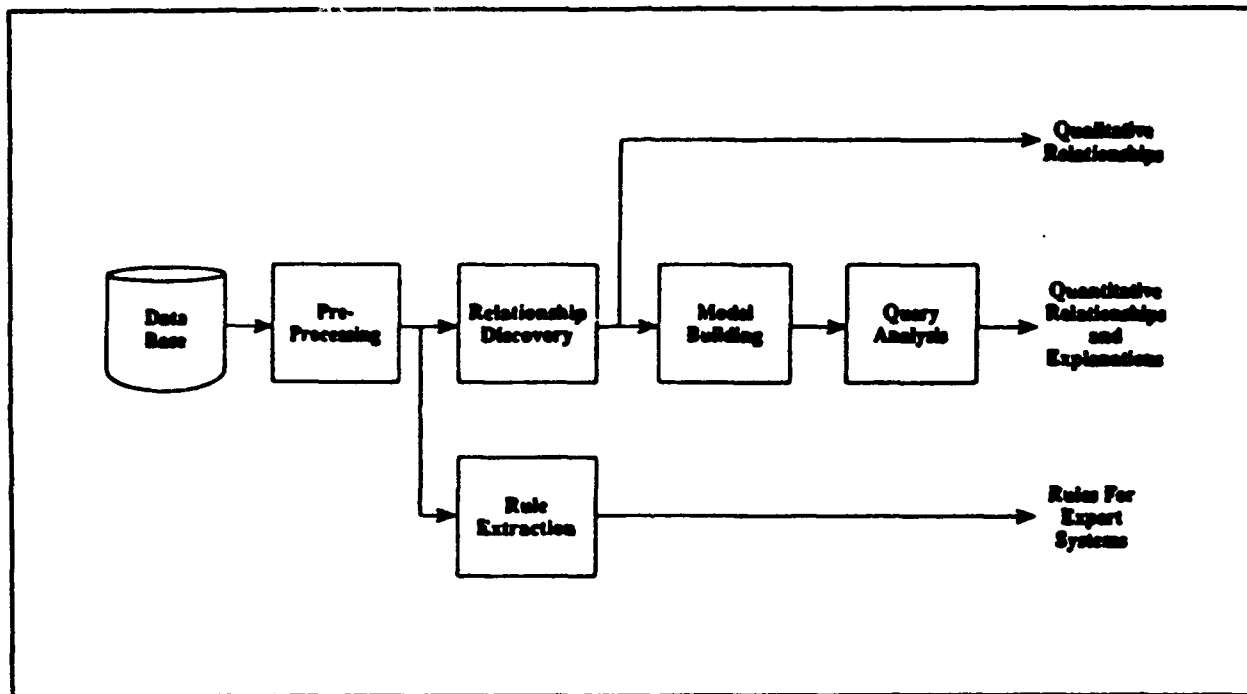


Figure 1. KENN Functional Data Flow

The KENN concept recognizes that the analysis is best accomplished in an incremental fashion. As such, KENN consists of five key analysis components. These components, shown in Figure 1, are:

- Preprocessing
- Relationship Discovery
- Model Building
- Query Analysis and Explanation
- Rule Extraction

The ability to automatically discover relationships contained within data, quantify their strength, and present them graphically to the user for visualization is defined as "Relationship Discovery". This capability was the major research effort during Phase I of this SBIR Project.

This Phase I SBIR project covered the enhancement and automation of the Relationship Discovery component. As seen in Figure 1, Relationship Discovery provides:

- Qualitative information on the existence of relationships between variables in the input data
- Input to the modeling component containing the information necessary to build smaller, more efficient models using subsets of the input variables.

Relationship Discovery in effect determines "what is important" in the data and estimates the strength of the relationships between the variables. This detection of relationships is the necessary precursor to the modeling step where the detected relationships are modeled.

The approach uses a self organizing neural network technique to approximate the probability density function (PDF) of the data set. This approximation of the PDF is necessary on large data sets to reduce the size of the problem and make it computationally tractable. The two-dimensional projections of the PDF are then automatically examined for the existence of relationships and a relationship strength value is assigned to the projection. The relationships with the highest strength are presented either rapidly as a two-dimensional scatter plot or as a three-dimensional graph of the PDF, with the Z dimension being the amplitude of the PDF at each point.

This approach is similar to correlation analysis, but provides more accurate and useful results since it can accommodate nonlinear as well as linear relationships. Additionally, this approach can handle relationships that are discontinuous or exist only over a limited range of values. Unlike conventional statistical approaches based on correlation analysis, the proposed Relationship Discovery technique makes no assumptions about the nature of any functional or nonfunctional relationships contained within the data. Instead, the capability is driven by the underlying probability density of the data, thereby giving a true representation of any relationships.

The Phase I effort for this SBIR concentrated on the development of the automated relationship discovery and relationship aggregation techniques. The objectives that drove this Phase I SBIR project effort are listed below:

- Refine and expand the analysis techniques used for the determination of the existence of a relationship between specific variables.
- Automate the relationship determination techniques.
- Quantify the "strength" of relationships and develop a mechanism to ignore relationships that have less than an automatically determined cutoff strength.
- Develop an automated process to aggregate the relationships and build relationship hierarchies.
- Develop proof-of-concept code that implements the above techniques.
- Perform tests of the code on simulated and real data.

Introduction to Relationship Discovery

The problem of Relationship Discovery can be phrased as follows: "Given a database, find subsets of the fields of the database such that knowing some of the fields in the subset provides information about other fields in the subset."

The assumption is that if A and B are related, then knowing A should provide some information about B. This relationship between A and B might then be exploited in some way. In particular, if one was building a model using A and B in the input set, one might be able to drop one of A and B from the input set because of the known relationship between the two.

The first observation to make about the problem stated above is that given a set of K variables, there are 2^K (two to the power of K) subsets of the K variables. Thus, it would be computationally infeasible to attempt to examine each of the subsets and determine a measure of the information content of the variables in the subset. Even if it were possible to do this, it would be unwieldy, to say the least, to attempt to use this huge mass of information.

Thus, we rephrase the problem in a slightly different way: "Given a database, determine a measure of the strength of the relationship between pairs of its fields."

A solution to this problem would provide a solution to any instance of the first problem for a specific subset of the variables. Given the subset one was concerned about, one would examine all pairs of its variables. If a pair of variables were found to be strongly related, it would be known that each one carries information about the other.

If we could solve this problem of quantifying the strength of relationships between the variables, we could then specify a cutoff strength below which relationships would be deemed insignificant. This would solve the problem of finding all relationships among the variables of the database.

What remains is to provide the modeling tool a way of aggregating the information contained in the relationship strengths to determine a reduced model to predict any given variable. In this context we note that related variables exhibit a certain kind of transitivity. If A is related to B and B is related to C, then A is related to C, but more weakly. This observation, phrased in terms of a distance measure that is inversely related to Relationship Strength, forms the basis of the Relationship Aggregation Component. Using this technique we build a reduced model by excluding variables that are linked indirectly to the variable we wish the model to predict.

In the following pages, we will describe approaches to reducing the problem size down to more manageable levels. We will then describe two approaches to the Relationship Discovery problem which quantify the strength of the relationships between variables in the input data. Finally, we will discuss the Relationship Aggregation component.

Throughout this report we will use the terms "field" (as in a database) and "variable" (as in a data set) interchangeably. We will also use the terms "record" (as in a database), "observation" (as in a data set), and "data point" (as in an n-dimensional space) interchangeably.

Decreasing the Size of the Problem

The time taken by the Relationship Discovery algorithms that will be described in the following sections goes up rapidly as the number of data points goes up. Often one will encounter problems in which there are more data points than one can process in a reasonable amount of time. In these cases, it is important to decrease the number of data points that will actually be used by the Relationship Discovery algorithms. There are several approaches to this.

The simplest approach is simply to sample the data. One can easily select a random subset of the observations that is as small as necessary. However, as the sample size grows smaller, the accuracy of the results gets worse. For our purposes, the accuracy would be unacceptably low.

One can do better if one chooses the sample with a little more effort. The k-means algorithm is an algorithm which generates representative data points that are cluster centers in the input data. This is accurately positions the points in the K-dimensional space, but does not accurately capture the relative probability density at different regions in the space. This is because clusters of many different sizes are all represented by a single cluster center each.

The Kohonen algorithm [5] with HNC's proprietary conscience mechanism provides a way of obtaining cluster centers that correctly represent the probability density function in K-space. This algorithm generates a set of points in the K-dimensional space spanned by the input such that each data point represents a specified fraction of the input data and is located close to the centroid of the data points it represents. This performs the kind of data compression needed by the Relationship Discovery algorithms.

Figures 2 and 3 show an example of the data compression provided the Kohonen algorithm. Figure 2 shows a set of input data points in a two dimensional space. Each input data point is represented by the symbol "x". X and Y are the variables spanning this space. In Figure 3, the same input data points are shown along with the cluster centers computed by the Kohonen algorithm with the conscience mechanism. Each cluster center is represented by a large dot connected by straight lines to the input data points it represents. It can be seen that the cluster centers are at the centroids of the clusters, and that each cluster center represents the same fraction (20%) of the input data points.

An interesting approach to decreasing the size of the Relationship Discovery problem involves the use of a hierarchical cluster tree. This is a tree of clusters in the input data having the following structure. At the top is the root, a single cluster containing all the input data points. The next level of the tree (children of the root) consists of a small number of super-clusters such that each data point belongs to exactly one super-cluster. At the next level (grandchildren of the root) there are a small number of clusters parented by each super-cluster. This hierarchical structure continues until at the bottom level, each data point is its own sub-sub-...-sub-cluster. This is illustrated in Figure 4.

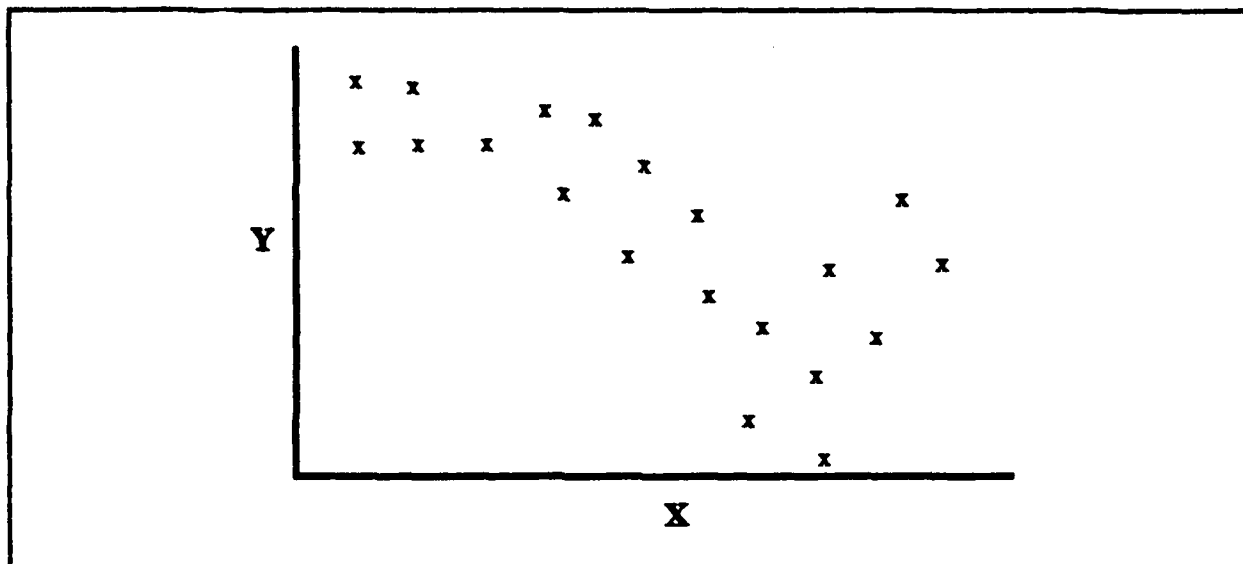


Figure 2: Raw Input Data

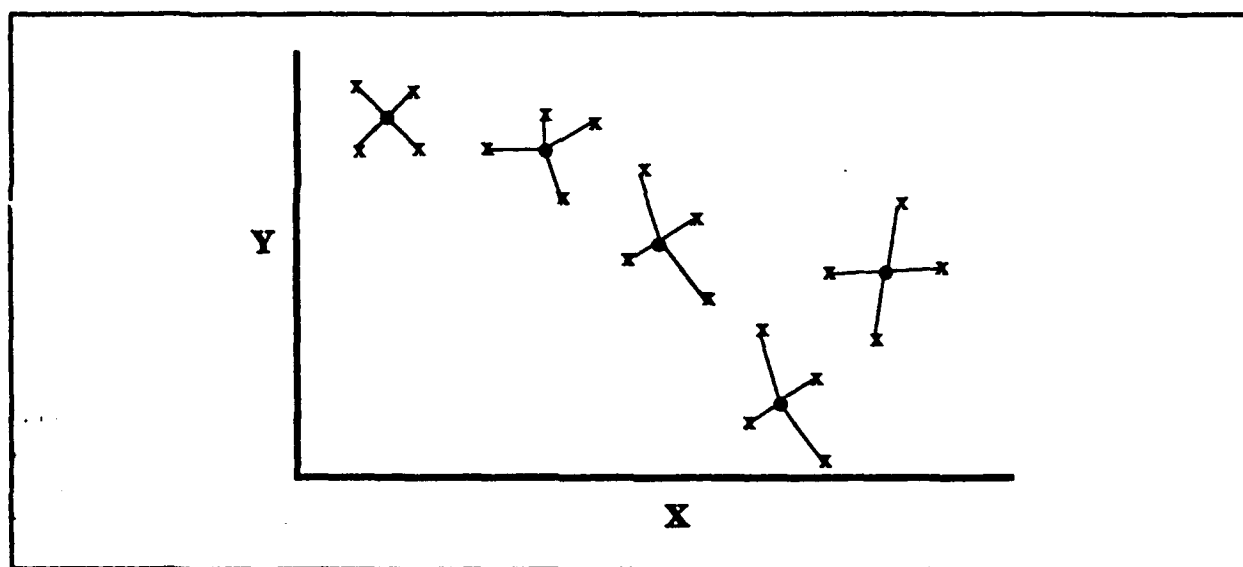


Figure 3: Data Compression using Kohonen Network

This kind of cluster tree could be built using the Kohonen algorithm with a form of node growth. Once it is built, the idea is to run Relationship Discovery on the super-clusters, clusters, sub-clusters, and so on until the relationships don't change much upon moving to the next level. This could potentially increase the efficiency of Relationship Discovery significantly.

Thus if very large problems are encountered, it is possible to decrease their size to more tractable levels. The Kohonen algorithm has been implemented and tested. The implementation of this hierarchical cluster tree concept was not within the scope of Phase I. It was also not proposed for implementation for Phase II as it is believed that the Kohonen approach to data compression would suffice for most real problems.

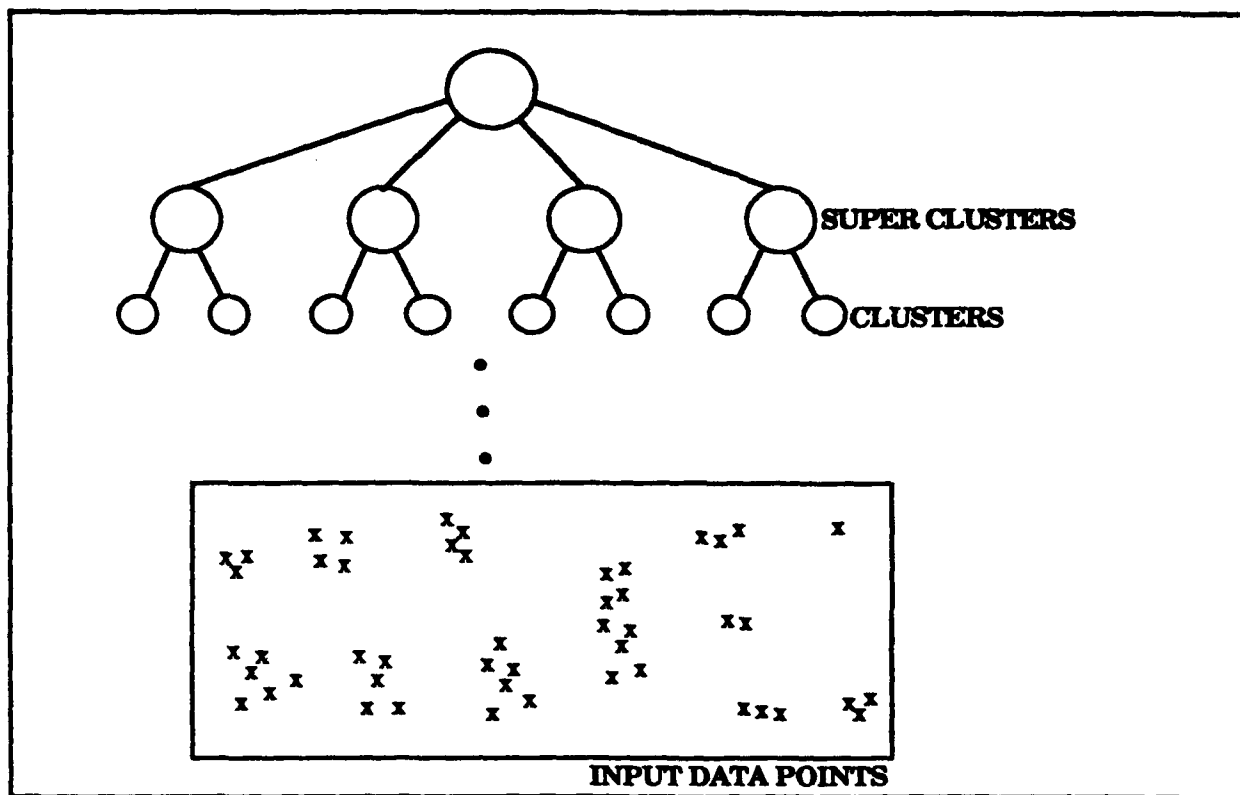


Figure 4: A Hierarchical Cluster Tree

Approaches to Relationship Discovery

Relationship Discovery Via the Chi-squared Test

Relationship Discovery using the Chi-squared test is performed as follows: Compress the data and form the projection of the compressed data set onto a pair of variables, say A and B. This gives us N ordered pairs of data points. Map these points into rank space. Bin the N ordered pairs into a small constant number of bins (a 5 by 5 grid of 25 bins is usually good enough). Compute the expected number of points in each bin, and the actual number of observations that fell into the bin. The Chi-squared statistic is a function of the number of data points that fell into each bin and the number that were expected to fall into each bin assuming that A and B were statistically independent. If the Chi-squared statistic is above a cutoff threshold for a given confidence level, there is a relationship. Do all this for all $O(K*K)$ non-redundant pairs of variables to determine all bivariate relationships among the variables.

Rank space for a particular variable is a space in which each value of the variable has been replaced by its rank in the set of all values of that variable. Thus, in rank space, the variable is uniformly (i.e., evenly) distributed.

The above approach to Relationship Discovery is based on the observation that if two variables have been transformed to rank space, they will be evenly distributed when considered individually, but they may or may not be evenly distributed when taken together. If they are evenly distributed when taken together (for example, when seen in a scatter plot), they are statistically independent. If they are not evenly distributed when taken together, we will say they are related.

Figure 5 shows a pair of variables, X and Y, in the original space and after being transformed to rank space. Note that once they are transformed to rank space, they are evenly distributed individually, but not when taken together. Consequently, X and Y are related.

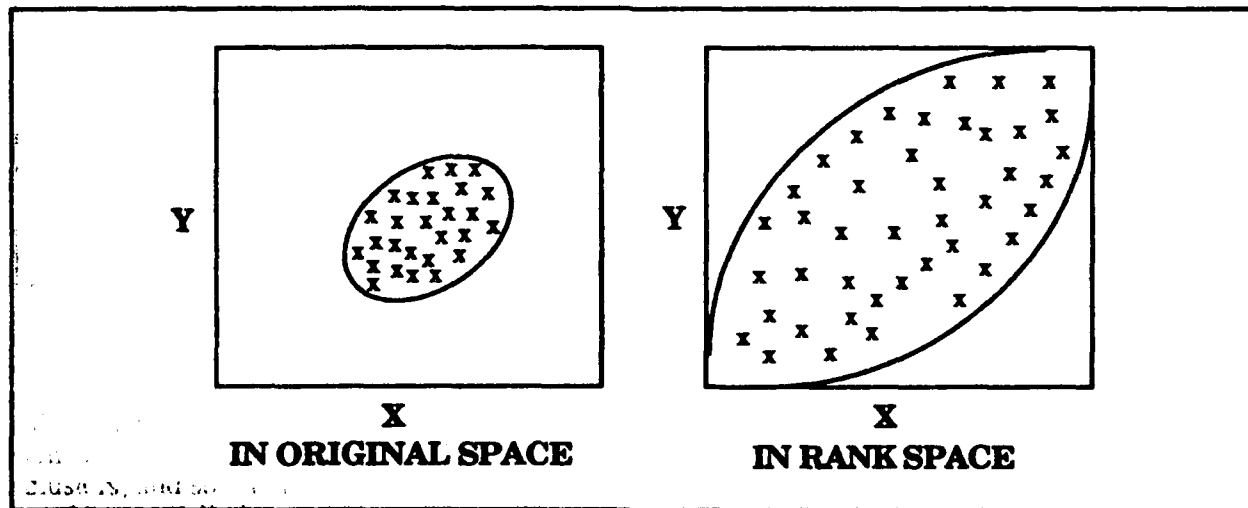


Figure 5: Transformation to Rank Space

Thus it very large project
level. The K-Nearest
neighbors cluster
algorithm
would suffice for this.

An illustration of the Chi-squared test is provided in Figure 6. Here the variables have been transformed to rank space and segmented into bins. All the bins do not have around the same number of data points in them, since there are areas that are too dense and too sparse. Because of the extent of the variation in the density of data points, the Chi-squared test deems there to be a relationship.

This approach is based on a strong theoretical foundation, since the Chi-squared test is a well-studied statistical test. It finds nonlinear as well as linear relationships. To find all bivariate relationships, this algorithm takes $O(K^2 * N)$ time. This time complexity makes the data compression operation crucial. Reducing N via data compression makes this algorithm feasible to execute.

This approach could be extended to find all relationships of third or higher order. It would be necessary to examine all distinct three-element subsets of the set of variables, except those containing two-elements subsets for which a relationship was already found. In the worst case, the complexity would be $O(K^3 * N)$. To find all m -variate relationships, the algorithm would take $O(K^m * N)$ time.

It is to be noted that the highest meaningful value of m is $O(\log N)$. For m greater than $O(\log N)$ there are just not enough data points to justify a claim that there is a m -variate relationship in the data.

This approach can be used for finding all low-order relationships in the data. This is reasonable since the relationships that are of greatest interest are of low order. Furthermore, low order relationships have more impact on the user since they can be displayed graphically for visualization by the user. Finally, almost all higher order relationships among finite statistical distributions have detectable lower order projections. Thus, if one detected all lower order relationships, one could infer the existence of almost all higher order relationships that would occur in practice with a Relationship Aggregation algorithm.

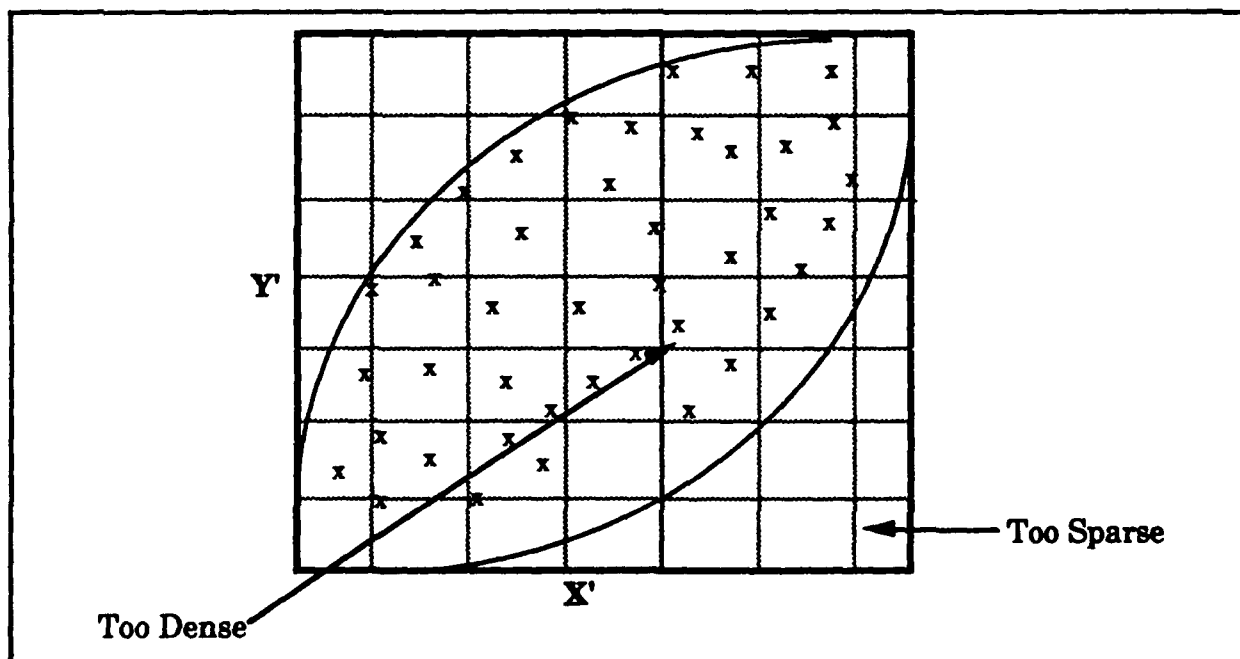


Figure 6: Relationship Discovery using the Chi-squared Test

Relationship Discovery Via Probing of K-space

The Probing of K-space approach to Relationship Discovery is performed as follows:

Transform the entire compressed data set (all variables) to a K-dimensional rank space. Thus each variable taken by itself has a uniform distribution. Take one of the input vectors, X , and perturb one of its components (say the i 'th component) to give X' . This is clearly illustrated in Figure 7. Find the nearest input vector to X' in a Euclidean K-space that has been stretched in the direction of the i 'th component. Let this nearest vector be Y . Call $(Y-X')$ the vector D .

The larger D is, the more sensitive the distribution is to the i 'th component (the component we initially perturbed). Most importantly, the components of D represent the sensitivity of all other dimensions to changes in the i 'th dimension.

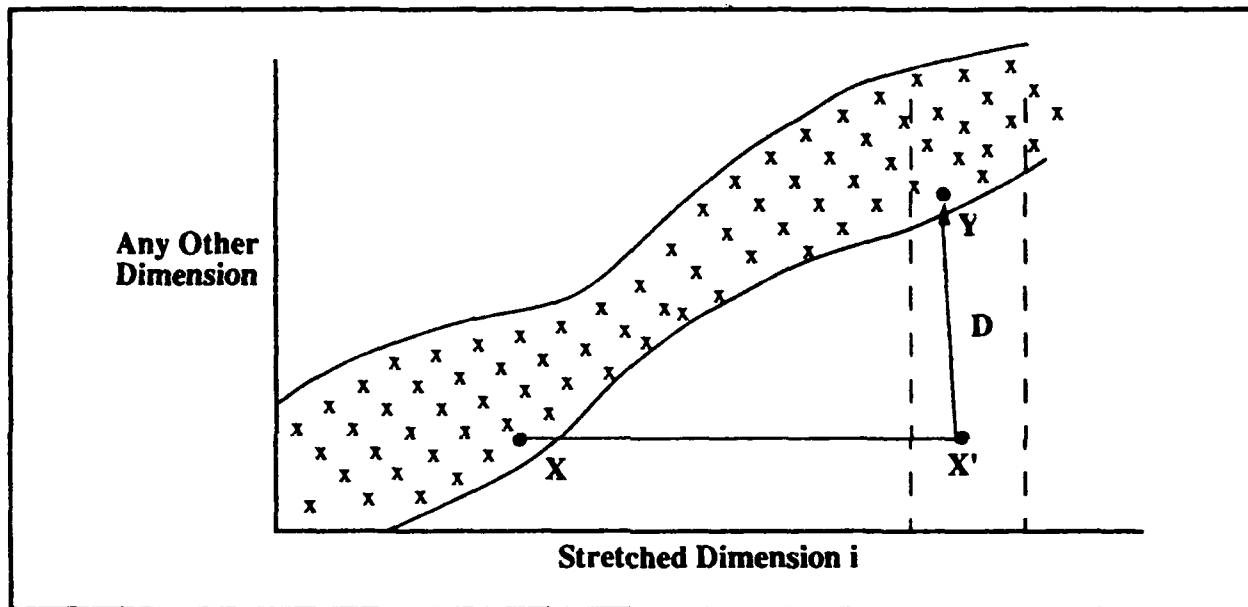


Figure 7: Probing of K-space

We are most interested in components of D that are consistently large in magnitude. So we compute this D vector using many different input vectors in turn as starting points. We end up with a distribution of the components of D , upon which we use the sum-of-squares statistic to measure whether a particular component is consistently large. A strong relationship is indicated by a high value of the sum-of-squares statistic.

We accumulate the sum-of-squares statistic for perturbations of each dimension, giving a square matrix that looks like a correlation matrix, but is different because it carries information about nonlinear relationships as well as linear ones.

This approach generates a square matrix of pairwise relationship strengths that contain information about both bivariate and higher order relationships. As was previously noted, the highest order relationship it can detect is $O(\log N)$, since there are not enough data points to justify a claim for the existence of relationships of higher order than $O(\log N)$.

This approach is a heuristic, and is similar in flavor to Sensitivity Analysis. Parameters that must be chosen well for this approach to work well are the stretch factor and the number of input data points to be perturbed for each dimension. Currently, we are obtaining good results with both the stretch factor and the number of input data points perturbed set to the square root of N . There is, however, no theoretical justification for this choice.

Relationship Aggregation

Both algorithms for Relationship Discovery described earlier provide quantitative measures of the strength of the relationships between pairs of the input variables. These strengths can be considered on an abstract level to be the weights of a weighted, undirected graph. Each variable is represented by a vertex and each relationship by an edge of this graph. Strongly related variables are linked by short edges and weakly related variables are linked by long edges. Totally unrelated variables are linked by infinitely long edges (or equivalently, nonexistent edges).

In this Relationship Graph, it is of interest to know the shortest path from one variable to another. That is, given two variables, through which other variables are they most strongly related? This kind of query will be addressed by the Relationship Aggregation Component.

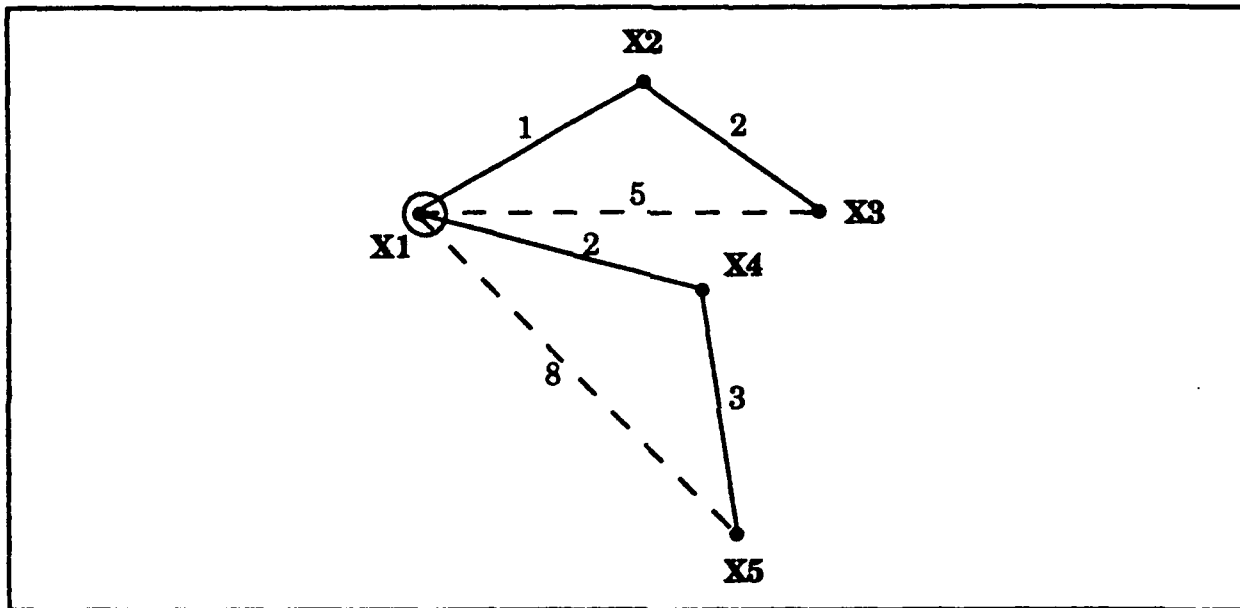
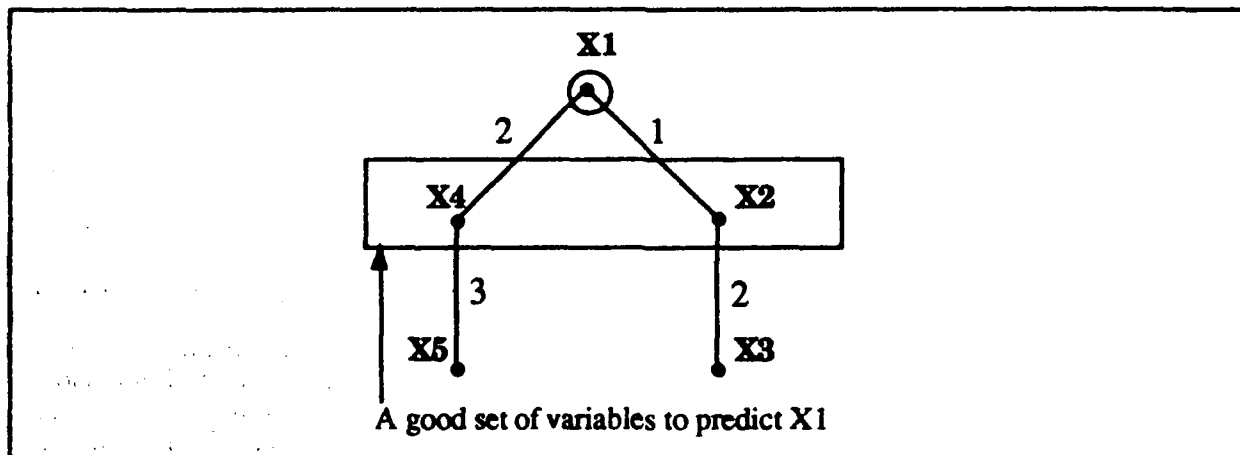


Figure 8a: A Relationship Graph and one of its Shortest Path Trees



Figures 8b: The Shortest Path Tree Shown from Figure 8a Separately

For example, one might have a Relationship Graph similar to the one shown in Figure 8a. In this figure, the nodes represent the variables in the data set, X1 through X5. The lines (both solid and broken) joining the nodes represent relationships among the variables. If two nodes are not connected then they are not related. The solid lines represent the relationships that are on the shortest path tree based on the variable X1.

In Figure 8b, the same shortest tree is shown in the format of a tree. The node labeled X1 is at the root of tree, and represents the variable X1.

The information contained in this tree representation is useful because it provides to the user structural information about the variables in the model.

Suppose the user was interested in indirect relationships between the variables X1 and X3. Upon querying the Relationship Aggregation system, the result would be: "The strongest relationship path from X1 to X3 is {X1,X2,X3}."

There are other methods to aggregate the information contained in the output of Relationship Discovery and condense it to present to the user. This approach has the advantage of having a clear visual interpretation as the variables are placed in a hierarchical tree structure. This allows the user to observe the strong linkages among the variables of the data set, organize them, and understand why variables are needed or not needed in the model.

Integrated Relationship Discovery Tool

The Integrated Relationship Discovery tool was built by merging the two forms of Relationship Discovery along with Relationship Aggregation together into a single software tool along with a data compression facility necessary to make the tool run efficiently.

This software tool is called "RELDISC" and works in tandem with a data preprocessing tool called SCALER. SCALER reads in a data file in ASCII format, scales the data to a range specified by the user, and writes the data to a binary file. RELDISC reads in this binary file, trains a Kohonen network, runs either or both Relationship Discovery algorithms, and performs Relationship Aggregation.

RELDISC allows the user to characterize the data set using self organization. This results in a fully trained Kohonen weight file representing the probability densities contained within the input database. This network, then, contains the "essence" of the database. Once the database has been characterized, RELDISC automatically searches for relationships and displays strong relationships. It then accepts user input to save particularly interesting PDF surfaces as data files that serve as input for off-line 3D plot programs. If desired, the user can then view selected relationships, perform relationship aggregation and generate minimum distance graphs.

RELDISC starts by training a Kohonen network to provide a compressed representation of the Probability Density Function of the input data. As this Kohonen net trains, its weights converge on points in the multi-dimensional input space that are the centers of equiprobable clusters. Thus, at the end of training, the weights of the Kohonen network provide a characterization of the entire Probability Density Function through expansion of the equiprobable cluster centers into Gaussians using the Parzen windowing process.

RELDISC allows the user to select one or both of the two Relationship Discovery algorithms described previously in this report. These algorithms have been implemented in C to run as a "stubnet" on the ANZA-Plus board.

It is to be noted that the Relationship Discovery algorithms are speeded up significantly by the data compression provided by the Kohonen network. The amount of processing that would be required if the data were not compressed in this manner would render the task infeasible for real-life sized data sets.

RELDISC feeds the output of Relationship Discovery to a viewing module that automatically displays as many of the strongest relationships as the user wishes to see. This can also be run in manual mode where the user requests to see a particular relationship.

RELDISC takes the results of Relationship Discovery and runs the Relationship Aggregation algorithm on them. This generates a "Shortest Path Tree" rooted at a user-specified variable. The user typically chooses to build the Shortest Path Tree from the variable which is to be the output of a backpropagation model. The tree contains information about the distances of other variables from the chosen variable, as well as information about variables that contain redundant information.

Once Relationship Aggregation is run, RELDISC generates a file containing linkage paths from user-chosen variables to the output variable. These paths are completely contained within the Shortest Path Tree, and consequently are Shortest Paths themselves.

Testing

Tests were performed on both approaches to Relationship Discovery. The test data consisted of artificially generated data from a linear congruential pseudorandom number generator. The data contained related and unrelated variables. Varying amounts of noise were added to the data, and the performance of Relationship Discovery was observed as a function of signal-to-noise ratio.

The following list contains the generating equations for a sample data set similar to those used for testing.

A = noise
B = A + noise
C = A*A + noise
D = cos(A) + noise
E = exp(-A*A) + noise
F = noise

The noise can be assumed to have a Gaussian distribution with mean 0 and variance 1.

The test results were as expected. As the amount of noise increased, the strength of the relationships detected (which represent the algorithm's confidence in its results) deteriorated. However, it is to be noted that the algorithms were able to take rather large amounts of noise before they erroneously ranked a non-relationship above a relationship.

Specifically, the PDF-above-Threshold algorithm performed correctly with a signal-to-noise ratio as low as 1.5:1, while the Probing-of-N-Space algorithm performed correctly with a signal-to-noise ratio as low as 2:1. (The performance was defined as correct if all the relationships had strengths that were greater than any of the non-relationships. In the above example, that would mean that all the relationships involving variable F would have lower strength than any of the others.)

Tests on Chi-squared Relationship Discovery

Test data set 1 was created using the equations shown in Table 1. The constant SN can be interpreted (in a loose way) as a signal-to-noise ratio. As SN becomes smaller, the amount of noise being added to the true functional relationship increases, as is evident in the equations. Test data set 2 is described by the equations in Table 2. The constant SN has the same interpretation. In both series of tests, the number of data points used was 1000.

<p>X = Gaussian random Linear = 2*X+noise/SN Quadratic = X*X + noise/SN Cosine = cos(X) + noise/SN Random = Gaussian random</p>

Table 1: Equations used to generate test data set 1

<p>X1 = Gaussian random</p> <p>X2 = Gaussian random</p> <p>Linear = $X1 + X2 + \text{noise/SN}$</p> <p>Quadratic = $X1 * X1 + X2 + \text{noise/SN}$</p> <p>Random = Gaussian random</p>

Table 2: Equations used to generate test data set 2

The results of Chi-squared Relationship Discovery on Test Data Set 1 are shown in Appendix 1. The results on Test Data Set 2 are shown in Appendix 2.

From examining the output of the tests, it can be seen that for low amounts of noise, the performance was 100% successful. However, as the amount of noise added to the functions increased, the performance of the algorithm deteriorated. It can be seen that a large amount of noise was added before the performance deteriorated significantly.

An important point to note is that in Test Data Set 2, the relationships were multivariate (higher order) and not just bivariate. However, the algorithm was able to detect the relationships between pairs of variables related through a higher order relationship. For example, the variables "X1" and "Linear" were found to be strongly related, as were "X2" and "Linear". It can be seen that this higher order relationship was detectable because of its projections. Putting the two relationships together, a Relationship Aggregation algorithm would be able to conclude that X1 and X2 were related, though less strongly. Thus the higher order relationship would be detected along with the bivariate (second order) relationships.

Tests on Probing-of-K-space Relationship Discovery

The results of Probing of K-space Relationship Discovery on Test Data Set 1 are shown in Appendix 3. The results on Test Data Set 2 are shown in Appendix 4.

From examining the output of the tests, it can be seen that for low amounts of noise, the performance was 100% successful from the point of view that all non-existent relationships were ranked below all relationships that existed. As the amount of noise added to the functions increased, the performance of the algorithm deteriorated. It can be seen that a large amount of noise was added before the performance deteriorated significantly. However, the performance was not as good as that of the Chi-squared test algorithm.

Tests using Civilian Customer Data

The Chi-squared Relationship Discovery algorithm was given further testing on a data set obtained from a civilian customer of HNC. This data set contained data on cellular phone usage and billing. The list of the relationships detected is given in Appendix 5. The relationships that were detected made intuitive sense, and upon looking at the data it could be seen that the relationships were actually present.

Extensions

Necessary Enhancements Identified in Phase I

During the attainment of these results and as a result of the testing of these concepts on data sets, several necessary enhancements to the overall Data Base Mining system concept were identified. These are:

- **Missing Value Prediction:** This capability will provide a way of systematically replacing missing values in the data set with maximum likelihood estimates of their true values.
- **Bad Data Detection:** This capability will help provide identification of potentially "bad" records in the data set.
- **Data Redundancy Removal:** This capability is necessary to optimize the model explanation capability and eliminate ambiguous results.

The first two of these, Missing Value Prediction and Bad Data Detection, together comprise Automatic Data Cleaning.

These enhancements identified during the Phase I effort are being proposed for implementation during Phase II. The technical approaches for these enhancements as well as their implications are detailed in this section.

Missing Value Prediction

Motivation for Missing Value Prediction

The problem of Missing Value Prediction is of significant practical importance in the analysis of data sets obtained from real-life databases. In typical databases, many records in the database contains at least one missing value. These missing values can significantly impact the perceived statistics of the database and adversely effect the resulting analysis. In building models, it is important to have good, clean data. The model can be no better than the data that is used to build the model. Missing values in the data set often constitute a significant problem in obtaining a good data set to build a high-quality model.

A formal definition of the missing value prediction problem can be stated as: "Given a database in which some of the variables for each of the observations are missing, determine good estimates for the missing values that are consistent with the other information in the same observation and in the rest of the data set."

Concept of Identity Map Network

To perform Missing Value Prediction, the Identity Map Network is used. The Identity Map Network is a Multi-layer Back-Propagation Network that is trained to reconstruct the input data at the output layer. As seen in Figure 9, the Identity Map Network has as many output nodes as it has inputs. However, the hidden layer has fewer nodes than either the input or the output layers. The concept of the identity map network was first described in [9].

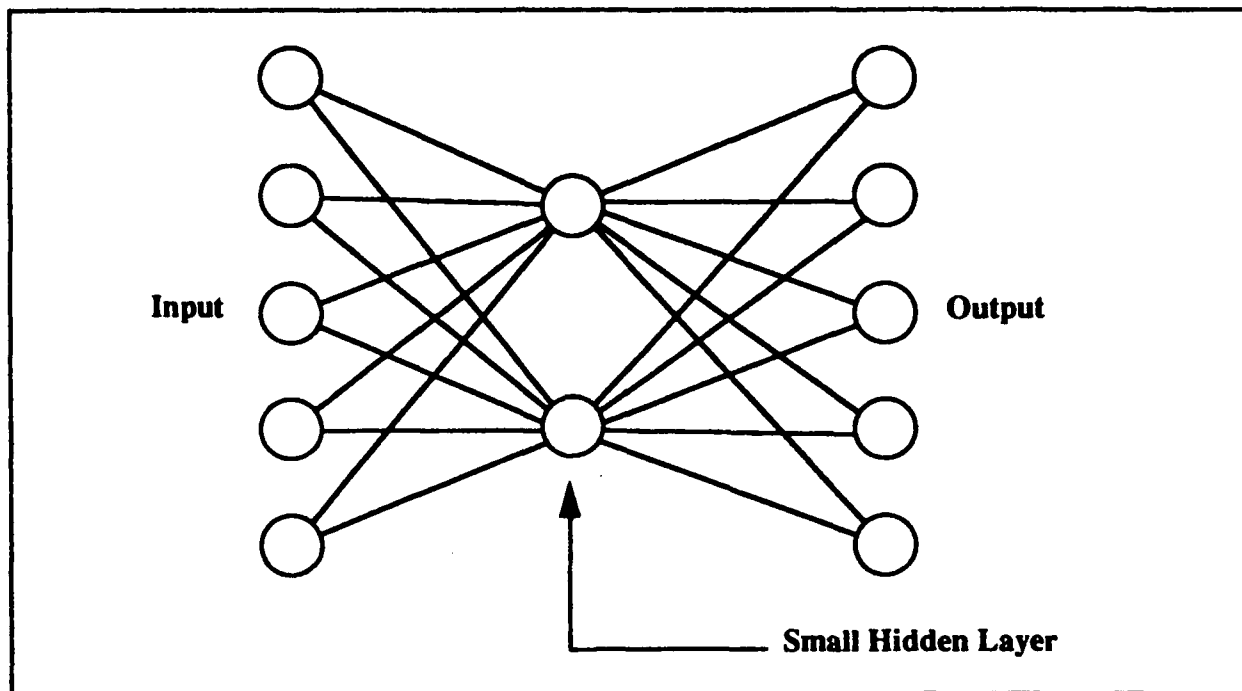


Figure 9: Identity Map Network

For simplicity of explanation, the solution of the problem is described using Identity Map Networks with only one hidden layer. However, it should be noted that Identity Map Networks with three hidden layers are significantly more powerful, and will likely be the architecture of choice for real-life problems. The techniques described here readily scale up to three hidden layer Identity Map Networks.

Considering the information flow through the Identity Map Network, it can be seen that the hidden layer acts as an "information bottleneck". As the network learns to reconstruct the input information at the output layer, it is forced to remove the redundancy from the original representation of the data and transmit only the high-information component of the original data through the hidden layer information bottleneck. It is also forced to learn to reconstruct the original data from the information transmitted through the bottleneck. Furthermore, in removing the redundancy in the data, the network learns the relationships between the fields in the data set.

Almost all data in real-life databases contains a significant amount of redundancy. The Identity Map Network is able to determine that redundancy and exploit it to solve the missing value prediction problem.

Relationships, Redundancy, and Data Manifolds

Consider a data set with k variables (equivalently, a database with k fields). Each of the k variables defines a dimension (coordinate) in a k -dimensional space. Each observation (or record) in this data set is a point in this k -dimensional space. If any of the variables are related to one another, there will be redundancy in the data set, since some variables contain information about others. Thus a relationship in the data is a redundancy. This redundancy will be exploited as part of the solution of the missing value prediction problem.

When the data set contains relationships, the data points will not occupy the entire k -dimensional space, but will lie on a lower-dimensional manifold in this space. The term "Data Manifold" is used to refer to the manifold on which the data points lie. As examples, a curve is a one-dimensional manifold while a surface is a two-dimensional manifold. In general it takes k values to specify a point in a k -dimensional space. However, if it is known that the point lies on a lower-dimensional manifold, say of p dimensions where $p < k$, it would take only p values to specify the point. Figure 10 shows an example where $k = 2$ and $p = 1$: a one-dimensional Data Manifold in a three dimensional space. In Figure 10, it can be seen that only one parameter is necessary to specify any point on the data manifold.

The Identity Map Network as a Data Manifold Approximator

If the data points lie on a p -dimensional manifold in a k -dimensional space (where $p < k$), then the architecture of the Identity Map Network would be k inputs nodes, p hidden nodes, and k output nodes. The information bottleneck would allow only p values to pass from the input side of the net to the output side in order to reconstruct the values of the k variables.

This network, in being trained to reconstruct the input at the output, is forced to learn the mapping from the original k -dimensional space to the p -dimensional Data Manifold in the first half of the network. This mapping to the p -dimensional manifold is represented by the activations on the p hidden units. In the second half of the network (from the hidden unit activation to the output), it is forced to learn the inverse mapping from the p -dimensional Data Manifold to the original k -dimensional space.

The activation values of the hidden nodes of this network constitute the values of p derived variables that parameterize the Data Manifold. Points on the Data Manifold are completely specified given values for these p variables. The second half of the Identity Map Network allows the translation of that specific manifold data point back into a point in the original k -dimensional space.

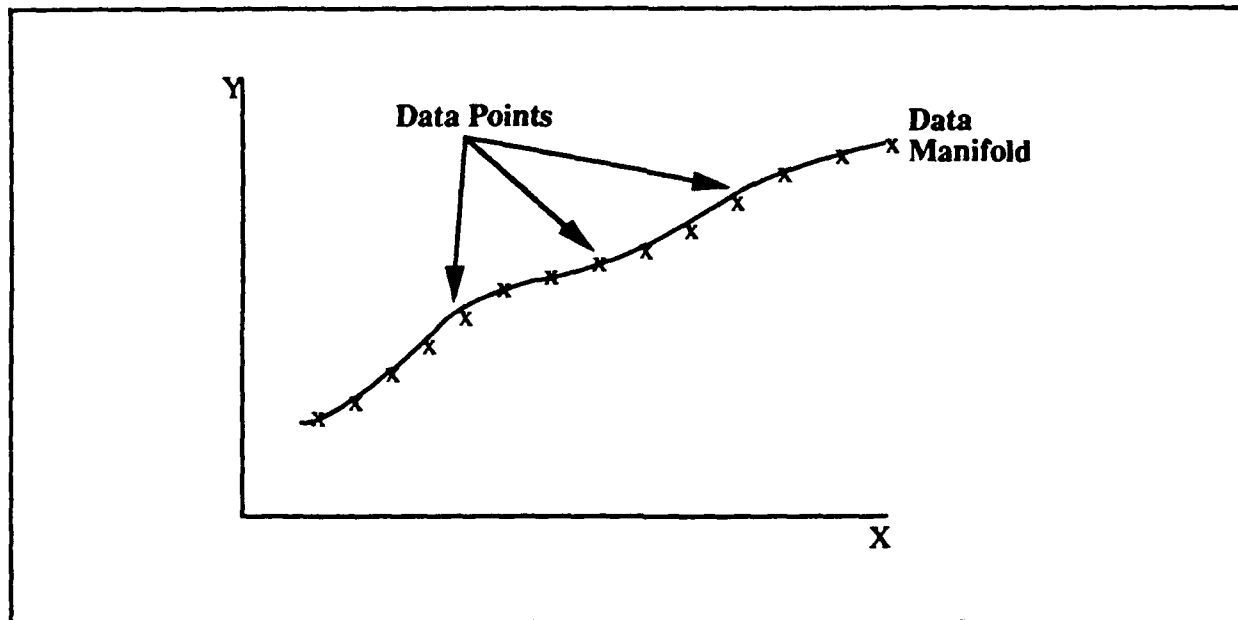


Figure 10: A One-dimensional Manifold in a Two-dimensional Space.

Determining the Dimension of the Data Manifold

In general, the dimension of the Data Manifold is not known *a priori*. The concept of the Automatic Neural Network (ANN) will be used to automatically determine the optimal net architecture, and thus the dimension of the Data Manifold. A description of the operation of the ANN for this problem is given below.

Assuming a k variable data set is given, the ANN will start with a net architecture that has k input nodes, one node in the hidden layer, and k output nodes. This network architecture will be trained using the ANN Back-Propagation Learning Algorithm, which will effectively perform gradient descent error minimization in the parameter space. In attempting to transmit the input information through the single-node bottleneck in the hidden layer, the network will learn the first (most important) "Generalized Principal Component" of the data. This will be the single derived variable that contains the most information for the reconstruction of the input data. Note that this variable is derived in a nonlinear way from the input data. Generalized Principal Components will be discussed in greater detail later.

Once the network converges (that is, continued learning only leads to small changes in the mean squared error that are below some threshold), learning on the weights associated with this part of the network is discontinued. A new node is then added in the hidden layer along with all its associated weights. These new weights are trained while keeping the weights associated with the previous node(s) frozen. This process of adding nodes is continued until an acceptable reconstruction quality is achieved. An example of a typical quality specification might be a mean squared reconstruction error of less than 0.0001 scaled units. The exact value of the error would, of course, be data set specific. When a reconstruction quality that meets or exceeds the requirements has been achieved, both the optimal network architecture and the dimension of the Data Manifold will have been determined. The dimension of the Data Manifold is the number of hidden nodes that were needed for good reconstruction.

Modifications to the Generic Backpropagation Learning Law

The Missing Values Problem brings with it two special requirements that necessitate minor modifications to the Back-Propagation learning law. Firstly, the obvious fact that the network must be trained with missing values raises a problem, since the Back-Propagation learning law makes no provisions for missing values in the data. The proposed solution to this problem is as follows:

- During forward propagation, the missing value is set to a random number. This is for the purpose of biasing the net towards not using the information in this variable. Each time this record is presented to the network, the missing value is set to a different random value.
- During backward propagation, the error for the missing value at the output node is set to zero, since nothing is known about this error.
- The second problem is due to a subtle point regarding learning with redundant information. Given a redundancy in the input data, the network has a large amount of choice regarding how it wishes to compress that redundancy and exactly what information it passes through the information bottleneck. For example:
 - Consider a two-variable set consisting of X and Y . Assume X and Y are linearly related. For simplicity consider a completely linear network. The network may choose to send the value of X , Y , or a linear combination of X and Y through the bottleneck in order to reconstruct both X and Y on the other side. The learning law will not bias it

either way since both contain identical information. Consequently the network will make a decision on this point that is predicated upon the initial conditions of the network, which are random.

- If neither value is missing, it doesn't matter which choice the network makes. However, suppose the network has chosen to transmit the value of X through the bottleneck and ignore the value of Y. Then a missing value of X would cause the network to be unable to reconstruct anything on the other side. This would make it impossible to predict missing X's from known Y's. However the opposite problem, that of predicting missing Y's from known X's, would cause the network no difficulty.
- Thus, when the network is given choices regarding what information to transmit through the bottleneck it would be preferable to have it choose to generate the transmitted information by blending the information from all its possible sources. This can be made to happen by penalizing the network during learning for taking too much information from any one source. Specifically this means adding to the error term for the network a penalty proportional to the sum of the squares of the weights in the pre-bottleneck layer. The network, in minimizing the total error, will attempt to drive down the values of large weights if this can be done without causing a deterioration in the reconstruction quality. These modifications to the learning law will result in the desired network behavior during training.

Predicting Missing Values

For the rest of this discussion, assume a trained Identity Map Network with an optimal architecture is available. The trained network will have discovered all the relationships and redundancy in the data. The network will be able to use this information to predict the missing values in the input data.

Assume there are k variables, and for some data point, only $k-1$ values for these variables are available, that is, one value is missing. Suppose variable i (where i is between 1 and k , inclusive) is the one whose value is unknown.

The following procedure is then used:

- Substitute a random value for the unknown variable (call it y_0) and cycle the data point through the trained network. This will provide a prediction of the unknown variable which will be called y_1 .
- Substitute y_1 as the i 'th variable in the data point, and cycle the data point again through the network.
- Continue this process, obtaining a sequence of predictions $\{y_0, y_1, y_2, \dots\}$ until the change in the predicted value between iterations approaches zero.

Figure 1 i shows a sequence of predictions for a typical case of Missing Value Prediction with a one dimensional manifold in a two dimensional space.

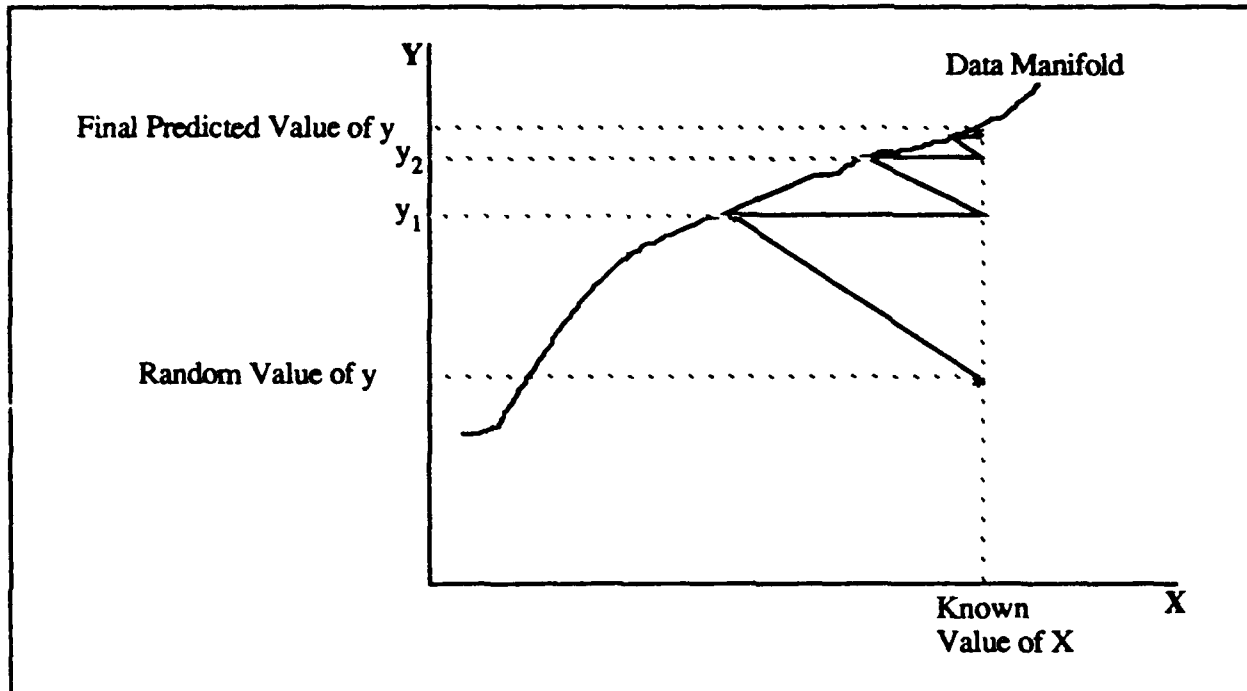


Figure 11: Missing Value Prediction

Note that all the known values are fed into the network without change at every iteration. Only the unknown value (which we are trying to predict) is being updated at every iteration with the output from the network at the previous iteration. This sequence has been observed in practice to always converge to four decimal places within 50 iterations. Convergence has been proved theoretically for the linear case. If the network was trained to a sufficient degree of accuracy, the sequence will converge to a value which is a good prediction of the variable on the basis of the values of the other variables in the data set.

If the variable with the missing value turned out to be one which did not have any redundant representations in the rest of the data set, then the sequence will converge immediately on the random value that was chosen initially. This is a rare case but is possible. To deal with this case where the data set does not contain the information needed to predict the missing variable, it would be appropriate to begin the iterations with the mode (the most probable value) of the missing variable across the data set, instead of a randomly chosen value. Then, even if no extra information is available, the predicted value will be the maximum likelihood estimate of the missing variable.

Predicting Multiple Missing Values in the Same Record

The same approach may be used to predict multiple missing values in the same record simultaneously. The known values are fed into the net at every iteration along with the latest prediction of the unknown values. The natural question that arises is that of the degradation of performance as the number of simultaneous missing values increases. That is, how many missing values can be tolerated simultaneously? The answer to this question lies in concepts of Generalized Principal Component Analysis, which will be discussed in a later section. Nevertheless a brief answer can be given as follows.

Missing values are predicted based on redundant information carried in other variables. If all the variables that could predict the missing variable are also missing, the prediction will not be accurate. But if some of the required variables are present to predict each of the missing values, the prediction will be accurate. In terms of Generalized Principal Component Analysis, the prediction of any particular variable will depend on the presence of at least one other variable along its Generalized Principal Component in the record.

Other Applications of This Technique

Another use of the Identity Map approach of modeling relationships and exploiting redundancy lies in the fact that a trained Identity Map Network contains within it the an estimate of the Generalized Principal Components of the input data. This approach of data compression can be used to reduce the size of the input variable set for modeling without losing valuable modeling information in the process. This will be discussed in detail in the section entitled "Generalized Principal Components".

Summary of Missing Value Prediction

The approach outlined above for solution to the missing value prediction problem is a novel technique that can effectively address the "cleaning" of large databases. Additionally, since this technique is closely coupled with the solution to the Data Redundancy Removal problem, these components and the associated reduced dimensionality feature vectors are available as a collateral advantage.

Bad Data Detection

Motivation for Bad Data Detection

Databases often contain incorrect data. This occurs for many reasons, including typographical errors during data entry and misinterpretation of the meaning of database fields. If left in the training set for the neural network, these "bad records" can significantly degrade the performance of the model. As mentioned before, the performance of the model is only as good as the data from which it is built. Eliminating erroneous records in the database will significantly improve model performance.

A formal definition of the bad record identification problem can be stated as: "Given a data set containing some bad records, determine a subset of the records which have fewer than a specified percentage of bad records. Furthermore, given a specific record, classify it as good or bad along with an estimate of the confidence of classification."

The following sections detail the proposed approach to solving the Bad Data Detection problem. The discussion in these sections first defines what is meant by "Bad Data". It then continues with a discussion of Robust Estimation. It proceeds with a description of how to obtain a robust variation of the Back-Propagation Learning Law. Finally, it describes how a Robust Identity Map Network can be used to solve the Bad Data Detection problem.

Definition of Bad Data

It is important to define what is meant by the term "bad data" in this context. It is not possible to look at a piece of data and state definitively whether it is bad or not. Nevertheless, it is feasible to determine whether it is a common or uncommon observation. That is, one can determine if there are many other similar observations in the data set or if this is a unique observation. The term "unique" is used loosely to mean that there are no observations close

to the data point under consideration. A more precise (but perhaps less intuitive) term for a unique data point would be an outlier, a data point that lies far from the mainstream.

If an observation is unique in the data set, it may be an error or it may be a rare but correct observation. If it is an error, it should be eliminated. If it is rare but correct, it should still be eliminated because conclusions made on small numbers of observations are not statistically valid. Thus it is fair to say that observations that are unique or occur very infrequently should be dropped from the data set. Such observations are termed "bad data".

Concept of Robust Estimation

Standard Back-Propagation as a Least-squares Estimator

The standard Back-Propagation Learning Law for neural networks performs gradient descent, iteratively improving the network weights to minimize the squared error of the network across the training set. Since the function being minimized is the squared error, the optimality condition is the least-squares criterion. Optimization with respect to the least-squares criterion provides maximum likelihood estimates of the neural network weights in the case that the error has a Gaussian distribution. Bad records in the database are problematic to the modeling process if they create outliers, which are data points lying far from the mainstream. If there are outliers in the data, the error in the data will not have a Gaussian distribution. Thus, if the optimality condition is the least-squares criterion, outliers will cause the network to deviate from the maximum likelihood estimates of the weights.

Gaussian Distribution of Error

An approach to solving this problem lies in using something other than the least-squares criterion as the optimality condition. Specifically, do not assume the errors to have a Gaussian distribution. With a Gaussian distribution of error, the probability of large errors decreases super-exponentially with the absolute value of the error. Thus if the network detects a large error, it decides it to be more probable that its weights are wrong than that the data point is wrong (since the probability that the data point is so wrong is super-exponentially small). Hence it makes a large correction in its weights in the direction of the data point with the large error. The correction in the weights is proportional to the size of the error. As the size of the error goes to infinity, the network's response increases without bound. In a sense, the network acts "gullible" in the presence of outliers when using the standard learning law because it blindly believes that every piece of data it sees is correct.

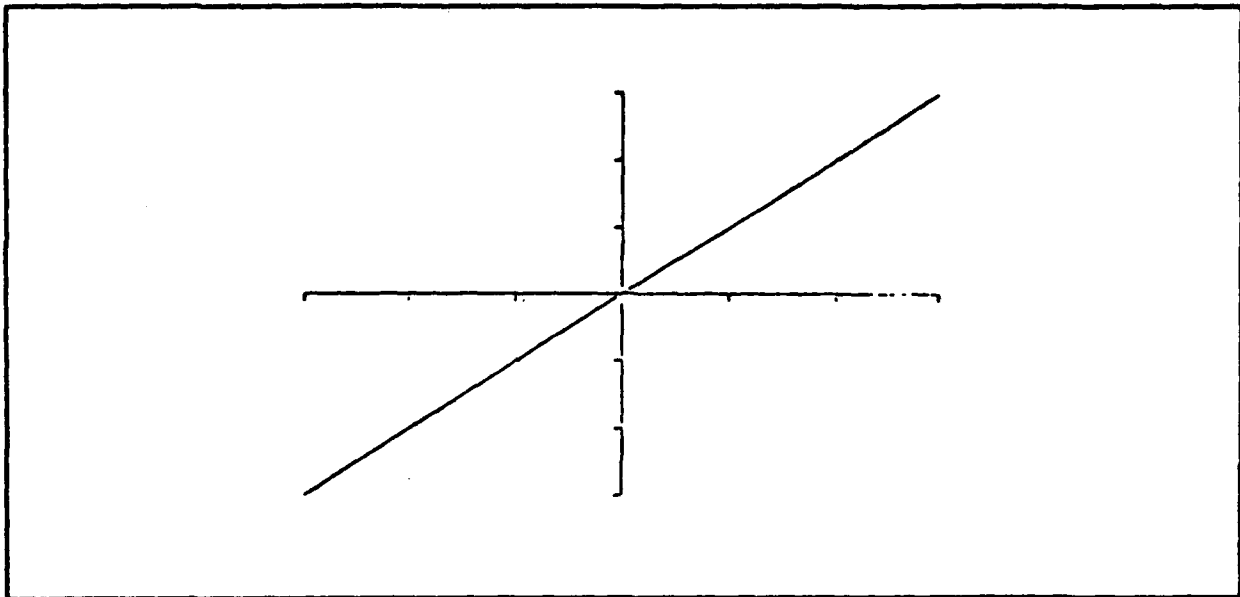


Figure 12: Response of the Network vs. Size of the Error Assuming a Gaussian Distribution

Cauchy Distribution of Error

If the error is assumed to have a fat-tailed distribution, such as a Cauchy distribution, the network will behave differently. If it encounters a data point that causes a large error, it decides that the data point is more likely to be wrong than right. Hence it makes practically no correction in its weights when it encounters large errors. This leads to behavior that can be considered "skeptical" in the presence of outliers. However, this kind of behavior can sometimes prevent the network from learning correctly at all.

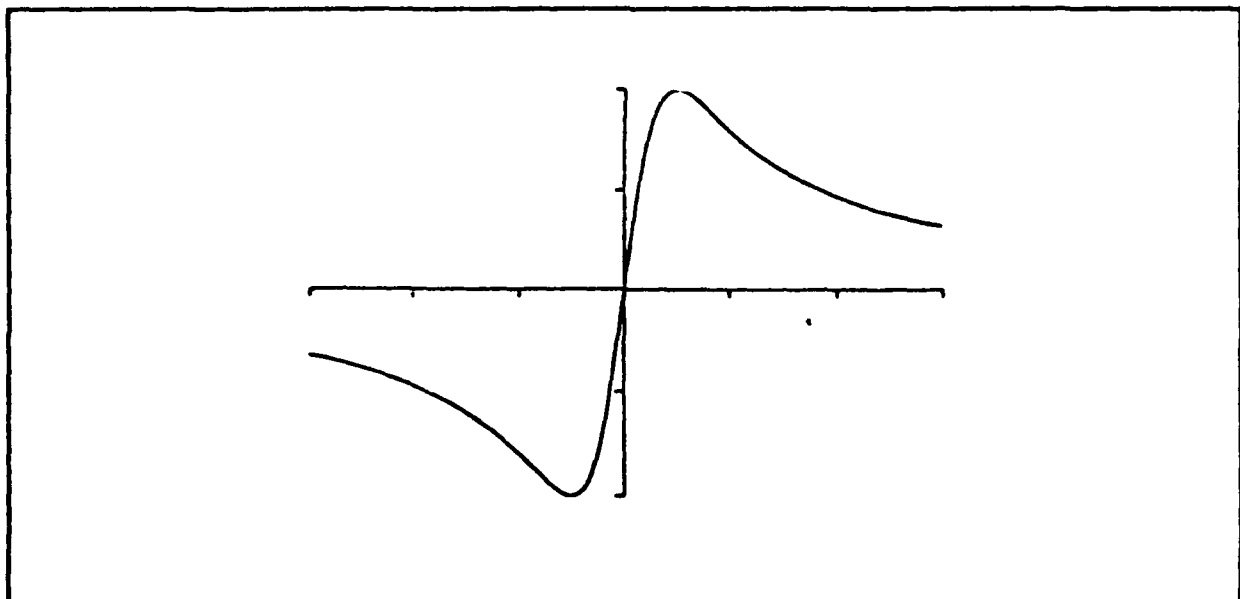


Figure 13: Response of Network vs. Size of the Error Assuming a Cauchy Distribution

Logistic Distribution of Error

A compromise between the above two behaviors can be obtained by assuming the error to have a Logistic distribution. With a learning law based on this assumption, the network will react similarly to moderate as well as large errors. Hence this kind of behavior can be considered to be neither "gullible" nor "skeptical", but "open-minded" in the presence of outliers. Using this approach, the network will be influenced by all points, but will gravitate towards those that appear consistently.

Assuming a Logistic distribution of error, the reaction of the network to large errors will not be proportional to the error, but to the hyperbolic tangent of the error. The hyperbolic tangent function lies in the range -1 to 1 everywhere and approaches the limits -1 and +1 at minus and plus infinity respectively. Thus no error, however large, will cause the network to be influenced by more than a constant amount.

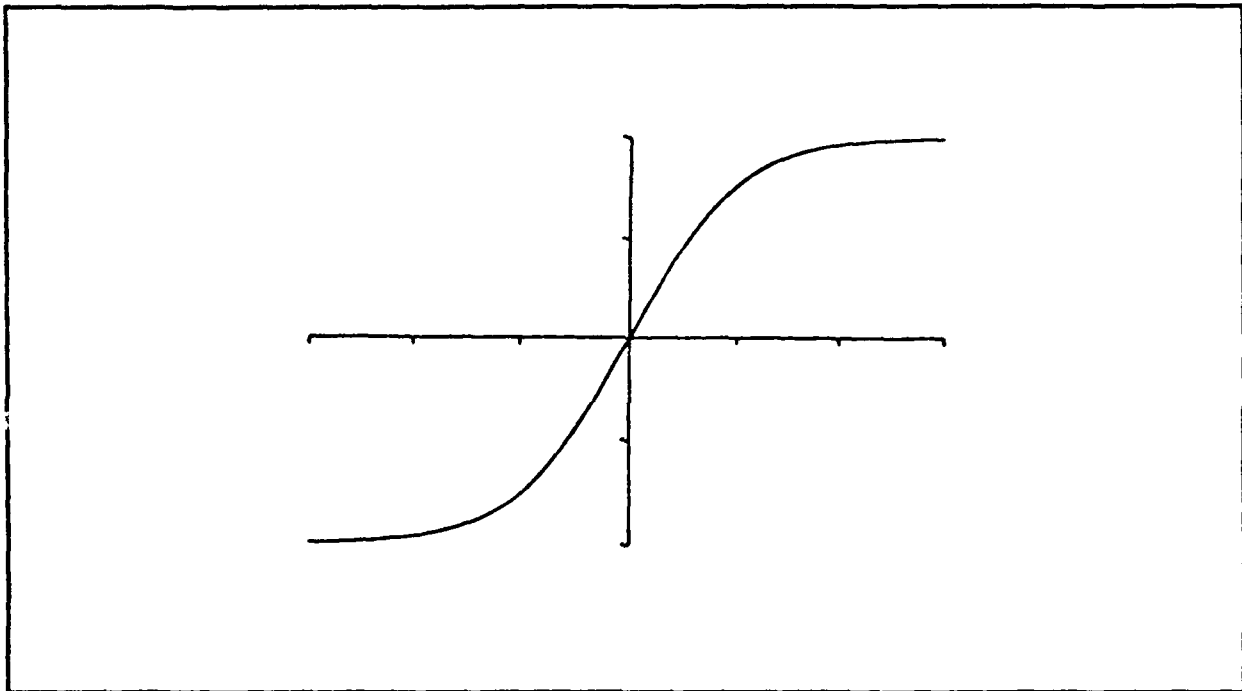


Figure 14: Response of the Network vs. Size of the Error Assuming a Logistic Distribution

This is the recommended modification to the Back-Propagation learning law for data containing outliers. In the following section this is used to perform robust estimation of the Data Manifold. The term "Robust Back-Propagation" is used to refer to Back-Propagation with the hyperbolic tangent function applied to the error term.

It should be noted that there is an expected tradeoff in the speed of convergence of Robust Back-Propagation in exchange for the property of robustness. This is due to the fact that the weights will not take large steps in response to large errors, and thus will need more iterations with smaller steps at each iteration in order to converge the same amount. Given that the ADAM Workstation will be based on the high throughput SNAP hardware, it is not felt that this is a concern.

Robust Estimation of the Data Manifold

Previously, estimation of the Data Manifold using the Identity Map Network was described. That method used the standard Back-Propagation learning law to train the Identity Map Network. In order to determine a robust estimate of the Data Manifold, the only change necessary is the use of the Robust Back-Propagation learning law instead of the standard Back-Propagation learning law. This network is termed the "Robust Identity Map Network".

Using the Robust Identity Map Network for Bad Data Detection

Given the Robust Identity Map Network, an approach to the problem of Bad Data Detection is addressed as follows:

- Take a data point and feed it into the trained network.
- Cycle the net once in feed-forward mode and observe the output.
- Compute the error of reconstruction.
- If the error of reconstruction is above a threshold value declare the data point to be erroneous.
- Otherwise declare the data point to be acceptable.

The concept is that since the network has been trained to reconstruct points on the mainstream of the Data Manifold correctly, badly reconstructed points do not lie on the Data Manifold. This concept is in unison with the basic ideas of Robust Estimation, where a large error is treated as evidence that a data point may be an outlier, and thus bad data. Since the network was trained using Robust Back-Propagation, this approach would be valid even for the data on which the network was trained. That is, one could expect large errors from outliers in the training data, which would not be the case if ordinary Back-Propagation had been used.

The size of the error relative to the good/bad decision threshold provides a confidence measure for the classification decision. In order to build a subset of the records with fewer than a specified percentage of bad records, one would merely have to adjust the threshold appropriately and extract the records that were classified as "good" from the data set.

Summary of Bad Data Detection

This approach to the Bad Data Detection problem can be used to "clean" existing databases. Additionally, this modification to the basic backpropagation learning law could be applied directly to modeling in order to build better quality models.

Data Redundancy Removal

Motivation for Data Redundancy Removal

The problem of Data Redundancy Removal arises in model reduction, where it is desired to determine a significantly reduced subset of the variables of a data set that is capable of prediction performance comparable to the full variable set. Model reduction is of particular importance in situations where the cost of collecting and/or processing the data is high relative to the returns. Data Redundancy Removal is also essential where post-modeling explanation is involved, since valid explanation requires the input to the model to be non-redundant, ie. reduced. If there is redundancy in the input data, the network has a large amount of choice regarding how it wishes to build the model using that redundant information. For example:

- Consider a three-variable set consisting of X, Y, and Z. Assume that $X=Y=Z$. For simplicity consider a linear network model with Z as the output. In other words, the network is trying to predict Z on the basis of X and Y.
- The network may choose to predict Z using X, Y, or a linear combination of X and Y. The learning law will not bias it either way since both contain identical information. Consequently the network will make a random decision on this point.
- After network training is completed, the next phase is typically network explanation using Sensitivity Analysis for aggregated information or Knowledge-Net for record-specific explanation. Here, partial derivatives of the network output with respect to its input are computed.
- Conceptually, the explanation results give the importance of each input in determining the output. In this case, the network may have randomly chosen to predict Z using X, Y, or a linear combination of X and Y. In each of these cases, the explanation results will be different. In each case the explanation will be correct, but none will provide a complete explanation of the relationships between the variables.

The example above shows that redundancy in the input data can confuse the results of neural network explanation. The statistical technique of Principal Component Analysis is capable of removing redundancy caused by linear correlations in the data. However, there is further redundancy caused by nonlinear relationships in the data. For reliable operation of the explanation mechanism, this redundancy must be removed also. Data Redundancy Removal provides the solution to this problem.

More formally, the redundancy elimination problem can be stated as: "Given a data set consisting of k variables, determine a minimal subset of the variables that contains the information present in the entire variable set, and use it to reconstruct the entire variable set to within a certain mean squared error tolerance."

Concepts of Data Redundancy Removal

Data Redundancy Removal can be viewed as a generalization of the traditional statistical method of Principal Component Analysis. To lay the groundwork for Data Redundancy Removal, the concepts of Principal Component Analysis are described here.

Basic Concepts of Principal Component Analysis

Given a data set, one may compute its correlation matrix, a square matrix whose entries represent the strength of the linear relationships between the variables. Then, one may perform an eigen-analysis on this correlation matrix. The eigenvectors of the correlation matrix are termed the Principal Components of the data set. These Principal Components are derived variables that contain information about redundancy in the input data caused by linear correlation. It should be noted that these derived variables are related to the original variables through purely linear relationships.

The Principal Components are uncorrelated with respect to one another and thus form a (linearly) non-redundant set that could be used to completely reconstruct the input data. The eigenvalues of the correlation matrix provide a measure of the amount of reconstruction information content associated with each eigenvector. Specifically, the largest eigenvalues are associated with the variables having the most information for reconstruction.

In order to specify the input data in a non-redundant way, it would be possible to specify the derived variables associated with the eigenvectors having the largest eigenvalues and drop the rest of the derived variables. A specified reconstruction error tolerance would determine how many derived variables (Principal Components) to retain. No information beyond the acceptable error tolerance would be lost in this process.

One approach to generalizing Principal Component Analysis

Linear Principal Component Analysis is capable of removing redundancy caused by linear correlations in the data. Further redundancy can be removed by considering the nonlinear relationships in the data. This, however, must be done with care, as the following discussion illustrates.

Principal Component Analysis uses an eigen-analysis of the correlation matrix in order to determine a set of derived variables that provide a linearly non-redundant specification of the input data to within a certain tolerance. Relationship Discovery provides a nonlinear equivalent of the correlation matrix called the Relationship Matrix. A natural approach that suggests itself is to perform an eigen-analysis of the Relationship Matrix to determine a set of derived variables that are non-redundant with respect to linear as well as nonlinear relationships.

However, a subtle point arises here. When all the redundancy that is removed is caused by linear relationships in the data, the input data can be directly reconstructed from the non-redundant representation with a linear transformation. However, if the redundancy that is removed is caused by a nonlinear relationship in the data, one must be careful to ensure that the reconstruction is feasible. In particular, while all the linear functions represented in the correlation matrix are invertible, the nonlinear functions represented in the relationship matrix are often not invertible.

For example, it is possible to predict X^2 uniquely from X . However, it is not possible to predict X uniquely from X^2 . Thus, when variables are highly related it is not always possible to predict an arbitrarily chosen one from another. Hence, one must use an approach which guarantees that reconstruction of the original data from the reduced representation is possible. This approach is described below.

The Identity Map Network for Data Redundancy Removal

The Identity Map Network described earlier for Missing Value Prediction computes what may be termed the "Generalized Principal Components" of the data set. Furthermore, it guarantees that reconstruction is possible by actually performing the reconstruction. Thus the Identity Map approach is a viable way of performing Data Redundancy Removal. The details are described in the following section.

Data Redundancy Removal using the Identity Map Network

The Identity Map Network is built as follows:

- Assume a k -variable data set is given. Start with a net architecture having k input nodes, one node in the hidden layer, and k output nodes. Train this network using the Robust Back-Propagation Learning Algorithm.
- When the mean squared error of the network converges, stop training on the weights associated with this part of the network.

- Then add a new node in the hidden layer along with all its associated weights. Train the weights associated with this new node, while keeping the weights associated with the first node constant.
- Continue the process of adding new nodes until an acceptable reconstruction quality is achieved.

In attempting to transmit the input information through the information bottleneck in the hidden layer, the network will learn derived variables that:

- Can reconstruct the input to the desired accuracy.
- Do not carry redundant information.
- May be derived in a nonlinear way from the input data.

These are nothing more than the Generalized Principal Components of the data. Note that the derived variables are exactly the variables that parameterize the Data Manifold. However, though the Generalized Principal Components of the data have been obtained, the problem of Data Redundancy Removal has not been completely solved. The next section details the remaining problem and its the solution.

Motivation for Realigning Generalized Principal Components

In Principal Component Analysis, the use of derived variables is an inconvenience. It would be preferable to merely select a subset of the input variables (rather than a set of derived variables) that could be used to reconstruct all the rest of the variables. In some situations the cost of collecting the data is high, so there is even more incentive to find a *small subset of the variables* that carries almost all the information in the full set. Another important problem is the difficulty of assigning meaning to derived variables, which causes explanation quality to suffer. Thus, there is significant motivation to seek a method of selecting a subset of the input variables that carry the information content of the full set.

The term for finding this subset of the input variables from the Generalized Principal Components is "realigning" the Generalized Principal Components

Realigning Linear Principal Components

In Principal Component Analysis, this "Realignment" capability exists, and is termed "rotation" of the Principal Components. The term arises for the following reason. The Principal Components represent an orthogonal set of coordinate axes in the input space. They span almost all the data, though not all of the space. The Principal Components are not in general aligned along the original coordinate axes. Thus, they form a set of variables derived by a general linear transformation of the input variables, not merely a subset of the input variables. However, if the Principal Components are rotated within the space they span in such a way that each input variable has a nonzero projection on exactly one Principal Component, the result is a "derived" variable set that is actually a subset of the original variable set.

Realigning Generalized Principal Components

In the case of Generalized Principal Component Analysis, the objective is the same but the approach is different. A mere rotation cannot align the Generalized Principal Components with the input variables because of the nonlinear relationships embodied in the Generalized Principal Components. However, consider the following. The first half of the Identity Map

Network is performing a transformation that generates the derived variables, while the second half of the Identity Map Network is regenerating the input variables from the derived variables. Place a restriction on the first half of the network such that only a selection is allowed, and not a general transformation. Then, the Generalized Principal Components that the network generates will be necessarily a mere selection of the input variables. The constraints on the network necessary to achieve this are commonly used in Linear Programming problems, where they are termed the "Assignment Constraints". Here they are referred to as the "Selection Constraints".

Use of the "Selection Constraints" transforms the neural network learning process from an unconstrained optimization problem to a constrained optimization problem. There is a standard technique used in optimization to deal with this situation: convert the problem back to unconstrained optimization by phrasing the constraints as terms in the objective function. That is, make the Selection Constraints penalty terms in the Back-Propagation Error function.

The Solution to the Data Redundancy Removal Problem

Thus, by modifying the Back-Propagation learning law to penalize violation of the Selection Constraints for the Identity Map Network, the network can be made to select a small subset of its inputs to reconstruct the output. This provides a way of building a reduced set of input variables that do not contain redundancy. This is the solution to the Data Redundancy Removal problem.

Summary of Data Redundancy Removal

In conclusion, Data Redundancy Removal can be solved using the approach described. The technique is closely aligned with both Missing Value Prediction and Bad Data Detection. This offers significant leverage in the technical development cycle and significantly reduces risk.

Conclusions and Discussion

Results of Phase I

All of the Phase I technical objectives have been met and alternate approaches to the relationship discovery process have been developed. Furthermore, a significant amount of refinement on the relationship discovery tool has been performed. This has resulted in a fully functional capability that is currently in use at HNC. The Relationship Discovery software tool along with an executable copy of the HNC Data Base Mining software will be delivered to the Army (AIRMICS) along with this Final Report.

Chi-squared Relationship Discovery

An automated Relationship Discovery algorithm based on a variant of the Chi-squared test has been designed and implemented. The Chi-squared test is a well-known test of statistical independence. In Relationship Discovery, it is used on the projected Probability Density Function. This approach has been tested on artificial data containing known relationships, and has achieved a 100% success rate in determining whether a relationship exists or does not exist in test problems.

Relationship Discovery via the Chi-squared test is a theoretically rigorous statistically-based approach to the problem. It detects all bivariate (second order) relationships in the data, and detects the projections of most multivariate (higher order) relationships found in real-life data. These detected relationships can then be visualized since they are all of low order.

Probing-of-K-space Relationship Discovery

An alternate algorithm has also been designed and implemented. This is termed "Probing of K-space", and is based on the idea of determining which other variables change when one variable changes in the projected Probability Density Function. This technique has been tested on artificial data containing known relationships, and has been able to rank the strengths of relationships among the variables successfully. This approach requires the user to specify a cutoff strength value below which relationships will be deemed to be insignificant.

The Probing of K-space approach is an alternate heuristic approach to the Relationship Discovery problem. It is more susceptible to noise than the Chi-squared test approach, and also is more computation-intensive. However, it has been successful in ranking relationships in order of decreasing strength in test cases.

Testing

Both of the above-mentioned approaches have also been tested on real data obtained from civilian customers of HNC. The results in these cases have been positive, since the algorithms captured many expected and intuitively obvious relationships in the data, as well as some that were not as intuitively obvious. The strongest relationships that fell in the latter category were inspected visually and it was confirmed that they did exist.

Relationship Aggregation

An algorithm for Relationship Aggregation has been designed and implemented. This algorithm derives a measure of the "distance" between pairs of variables from the previously obtained relationship strengths, and treats them as weights in a complete, weighted, undirected graph. The algorithm then finds the shortest path between variables in this graph, builds a tree of such shortest paths, and presents this tree to the user. This tree captures the most relevant relationships in the data and allows the user to grasp them readily.

Relationship Aggregation using the Shortest Path approach is a heuristic, visually oriented method of helping the user to analyze the variable structure and determine a reduced model to predict a given variable. It is able to analyze the output of either of the two methods of Relationship Discovery.

Integrated Relationship Discovery Tool

An Integrated Relationship Discovery software tool has been built. This tool contains the two forms of Relationship Discovery, Relationship Aggregation and a Kohonen-based data compression facility necessary to make the Relationship Discovery algorithms run efficiently.

Necessary Enhancements Identified in Phase I

During the attainment of these results and as a result of the testing of these concepts on data sets, several necessary enhancements to the overall Data Base Mining system concept were identified. These are:

- **Missing Value Prediction:** This capability will provide a way of systematically replacing missing values in the data set with maximum likelihood estimates of their true values.
- **Bad Data Detection:** This capability will help provide identification of potentially "bad" records in the data set.
- **Data Redundancy Removal:** This capability is necessary to optimize the model explanation capability and eliminate ambiguous results.

The first two of these, Missing Value Prediction and Bad Data Detection, together comprise Automatic Data Cleaning.

These enhancements identified during the Phase I effort are being proposed for implementation during Phase II.

Summary of Results Obtained in Phase I

In summary, the following results were achieved during Phase I of this SBIR project:

- Implementation and demonstration of an automatic approach to relationship discovery via 2-D subspace projections.
- Development of a Chi-squared test to find relationships.
- Implementation of the capability for a scatter-plot display for quick visualization of relationships.

- Development of the capability to perform efficient Parzen windowing on the approximated PDF to generate 3-D surface plots of the PDF projections.
- Development of the capability to display relationships in rank space.
- Development of an approach to the model reduction problem via the shortest path method using the Chi-squared test and Cramer's coefficient to generate the Shortest Path Tree.
- Implementation of these capabilities in software and verification of software performance on real and artificial data.
- Enhancement of the general software capabilities of the Relationship Discovery tool:
 - Increase in utilization of neurocomputer processing capabilities to reduce processing time.
 - Enhancement of usability through implementation of a simple graphical user interface.
 - Integration of the software into a single coalesced environment.
- Conceptualization of approaches to three necessary enhancements to the Data Base Mining Concept.

Technical References

1. Hecht-Nielsen, R. "Kolmogorov's Mapping Neural Network Existence Theorem", Proceedings First IEEE Conference on Neural Networks, Vol. II, pp 11-14, San Diego, 1987.
2. Hecht-Nielsen, R. "Theory of Backpropagation Neural Networks", Proceedings First Joint Conference on Neural Networks, Vol. I, pp 593-608, San Diego, 1988.
3. White, H., et al, "Multilayer Feed forward Networks are Universal Approximators", Neural Networks, Vol. 2, No. 5, pp 359-366, 1989.
4. White, H. Stinchcombe, M. "Universal Approximation Using Feed forward Networks with Non-sigmoid Hidden Layer Activation Functions", Proceedings First Joint Conference on Neural Networks, Vol. I, pp 613-618, San Diego, 1988.
5. Kohonen, T. "Self-Organization and Associative Memory", Second Edition, Springer-Verlag, 1988.
6. Parzen, E. "On Estimation of a Probability Density Function and Mode", Annals of Mathematical Statistics, Vol. 33, Sept. 1962.
7. Duda P., Hart, P. "Pattern Classification and Scene Analysis", Wiley, New York, 1973.
8. Fahlman, Scott E. "An Empirical Study of Learning Speed in Backpropagation Networks", Technical Report CMU-CS-88-162, Department of Computer Science, Carnegie-Mellon University, June 1988.
9. Rumelhart, David E., et al, "Parallel Distributed Processing", Volume 1. MIT Press, 1986.
10. White, Halbert, "Supervised Learning as Stochastic Approximation", International Joint Conference on Neural Networks, San Diego, CA., June 1990.

APPENDICES

Appendix 1

Chi-squared Relationship Discovery on Test Data Set 1

This appendix contains the output of Chi-squared Relationship Discovery run on Test Data Set 1. The equations used to generate this test data set are:

$X = \text{Gaussian random}$

$\text{Linear} = 2 * X + \text{noise} / \text{SN}$

$\text{Quadratic} = X * X + \text{noise} / \text{SN}$

$\text{Cosine} = \cos(X) + \text{noise} / \text{SN}$

$\text{Random} = \text{Gaussian random}$

As mentioned in the text, the constant SN is similar to a signal-to-noise ratio to the extent that small values of SN imply large amounts of noise being added to the function values.

Five different values of SN were used. The value of SN for each run is printed on the output page.

The number of data points used was 1000.

x --	linear :	76.97	RELATIONSHIP DETECTED
quadratic --	cosine :	40.86	RELATIONSHIP DETECTED
x --	quadratic :	30.03	RELATIONSHIP DETECTED
linear --	quadratic :	27.85	RELATIONSHIP DETECTED
x --	cosine :	24.63	RELATIONSHIP DETECTED
linear --	cosine :	22.30	RELATIONSHIP DETECTED
cosine --	random :	0.51	
quadratic --	random :	0.36	
x --	random :	0.28	
linear --	random :	0.19	

Chi-squared Relationship Discovery

Test Data Set 1

SN = 10

x --	linear :	56.16	RELATIONSHIP DETECTED
quadratic --	cosine :	25.24	RELATIONSHIP DETECTED
x --	quadratic :	22.33	RELATIONSHIP DETECTED
linear --	quadratic :	17.40	RELATIONSHIP DETECTED
x --	cosine :	16.91	RELATIONSHIP DETECTED
linear --	cosine :	13.16	RELATIONSHIP DETECTED
quadratic --	random :	0.41	
x --	random :	0.28	
cosine --	random :	0.22	
linear --	random :	0.15	

Chi-squared Relationship Discovery

Test Data Set 1

SN = 5

x --	linear :	27.48	RELATIONSHIP DETECTED
x --	quadratic :	14.44	RELATIONSHIP DETECTED
quadratic --	cosine :	9.10	RELATIONSHIP DETECTED
linear --	quadratic :	7.41	RELATIONSHIP DETECTED
x --	cosine :	7.03	RELATIONSHIP DETECTED
linear --	cosine :	3.62	RELATIONSHIP DETECTED
quadratic --	random :	0.42	
linear --	random :	0.42	
cosine --	random :	0.36	
x --	random :	0.28	

Chi-squared Relationship Discovery

Test Data Set 1

SN = 2

x --	linear :	12.70	RELATIONSHIP DETECTED
x --	quadratic :	7.59	RELATIONSHIP DETECTED
x --	cosine :	2.83	RELATIONSHIP DETECTED
quadratic --	cosine :	2.22	RELATIONSHIP DETECTED
linear --	quadratic :	1.59	RELATIONSHIP DETECTED
linear --	cosine :	0.67	
quadratic --	random :	0.48	
cosine --	random :	0.46	
linear --	random :	0.31	
x --	random :	0.28	

Chi-squared Relationship Discovery

Test Data Set 1

SN - 1

x --	linear :	3.95	RELATIONSHIP DETECTED
x --	quadratic :	3.16	RELATIONSHIP DETECTED
x --	cosine :	0.94	
quadratic --	cosine :	0.60	
quadratic --	random :	0.50	
cosine --	random :	0.39	
linear --	random :	0.29	
x --	random :	0.28	
linear --	cosine :	0.27	
linear --	quadratic :	0.18	

Chi-squared Relationship Discovery

Test Data Set 1

SN = 0.5

Appendix 2

Chi-squared Relationship Discovery on Test Data Set 2

This appendix contains the output of Chi-squared Relationship Discovery run on Test Data Set 2. The equations used to generate this test data set are:

$X1 = \text{Gaussian random}$

$X2 = \text{Gaussian random}$

$\text{Linear} = X1 + X2 + \text{noise}/\text{SN}$

$\text{Quadratic} = X1 * X1 + X2 + \text{noise}/\text{SN}$

$\text{Random} = \text{Gaussian random}$

As mentioned in the text, the constant SN is similar to a signal-to-noise ratio to the extent that small values of SN imply large amounts of noise being added to the function values.

Five different values of SN were used. The value of SN for each run is printed on the output page.

The number of data points used was 1000.

x2 --	quadratic :	23.75	RELATIONSHIP DETECTED
linear --	quadratic :	18.17	RELATIONSHIP DETECTED
x2 --	linear :	13.67	RELATIONSHIP DETECTED
x1 --	linear :	12.74	RELATIONSHIP DETECTED
x1 --	quadratic :	9.41	RELATIONSHIP DETECTED
linear --	random :	0.40	
x1 --	x2 :	0.40	
quadratic --	random :	0.38	
x1 --	random :	0.26	
x2 --	random :	0.09	

Chi-squared Relationship Discovery

Test Data Set 2

SN = 10

x2 --	quadratic :	22.47	RELATIONSHIP DETECTED
linear --	quadratic :	14.97	RELATIONSHIP DETECTED
x2 --	linear :	13.31	RELATIONSHIP DETECTED
x1 --	linear :	12.49	RELATIONSHIP DETECTED
x1 --	quadratic :	9.59	RELATIONSHIP DETECTED
linear --	random :	0.45	
x1 --	x2 :	0.40	
quadratic --	random :	0.39	
x1 --	random :	0.26	
x2 --	random :	0.09	

Chi-squared Relationship Discovery

Test Data Set 2

SN - 5

x2 --	quadratic :	16.10	RELATIONSHIP DETECTED
x2 --	linear :	12.03	RELATIONSHIP DETECTED
x1 --	linear :	10.64	RELATIONSHIP DETECTED
linear --	quadratic :	9.12	RELATIONSHIP DETECTED
x1 --	quadratic :	8.88	RELATIONSHIP DETECTED
linear --	random :	0.47	
x1 --	x2 :	0.40	
quadratic --	random :	0.29	
x1 --	random :	0.26	
x2 --	random :	0.09	

Chi-squared Relationship Discovery

Test Data Set 2

SN = 2

x2 --	quadratic :	9.17	RELATIONSHIP DETECTED
x2 --	linear :	8.14	RELATIONSHIP DETECTED
x1 --	linear :	7.26	RELATIONSHIP DETECTED
x1 --	quadratic :	7.08	RELATIONSHIP DETECTED
linear --	quadratic :	3.34	RELATIONSHIP DETECTED
linear --	random :	0.50	
x1 --	x2 :	0.40	
quadratic --	random :	0.30	
x1 --	random :	0.26	
x2 --	random :	0.09	

Chi-squared Relationship Discovery

Test Data Set 2

SN - 1

x2 --	linear :	4.02	RELATIONSHIP DETECTED
x1 --	quadratic :	3.90	RELATIONSHIP DETECTED
x2 --	quadratic :	3.89	RELATIONSHIP DETECTED
x1 --	linear :	3.25	RELATIONSHIP DETECTED
linear --	quadratic :	0.75	
quadratic --	random :	0.40	
x1 --	x2 :	0.40	
linear --	random :	0.32	
x1 --	random :	0.26	
x2 --	random :	0.09	

Chi-squared Relationship Discovery

Test Data Set 2

SN = 0.5

Appendix 3

Probing of K-space on Test Data Set 1

This appendix contains the output of Probing of K-space run on Test Data Set 1. The equations used to generate this test set are:

$X = \text{Gaussian random}$

$\text{Linear} = 2 * X + \text{noise} / \text{SN}$

$\text{Quadratic} = X * X + \text{noise} / \text{SN}$

$\text{Cosine} = \cos(X) + \text{noise} / \text{SN}$

$\text{Random} = \text{Gaussian random}$

As mentioned in the text, the constant SN is similar to a signal-to-noise ratio to the extent that small values of SN imply large amounts of noise being added to the function values.

Five different values of SN were used. The value of SN for each run is printed on the output page.

The number of data points used was 1000.

x --	linear :	99.98
x --	quadratic :	51.71
quadratic --	cosine :	33.18
linear --	quadratic :	31.95
x --	cosine :	30.70
linear --	cosine :	29.04
quadratic --	random :	24.60
linear --	random :	23.79
cosine --	random :	22.58
x --	random :	16.94

Probing of K-space Relationship Discovery

Test Data Set 1

SN = 2

x --	linear :	99.96
x --	quadratic :	73.32
x --	cosine :	65.57
linear --	cosine :	61.46
quadratic --	cosine :	58.79
linear --	quadratic :	51.98
quadratic --	random :	50.36
linear --	random :	49.58
cosine --	random :	48.12
x --	random :	46.82

Probing of K-space Relationship Discovery

Test Data Set 1

SN = 1

cosine --	random :	99.95
x --	linear :	93.01
x --	cosine :	90.32
linear --	random :	87.14
quadratic --	random :	86.53
x --	quadratic :	85.60
linear --	cosine :	80.96
quadratic --	cosine :	77.16
x --	random :	72.52
linear --	quadratic :	70.44

Probing of K-space Relationship Discovery

Test Data Set 1

SN = 0.5

Appendix 4

Probing of K-space on Test Data Set 2

This appendix contains the output of Probing of K-space run on Test Data Set 2. The equations used to generate this test data set are:

$X1 = \text{Gaussian random}$

$X2 = \text{Gaussian random}$

$\text{Linear} = X1 + X2 + \text{noise}/\text{SN}$

$\text{Quadratic} = X1 * X1 + X2 + \text{noise}/\text{SN}$

$\text{Random} = \text{Gaussian random}$

As mentioned in the text, the constant SN is similar to a signal-to-noise ratio to the extent that small values of SN imply large amounts of noise being added to the function values.

Five different values of SN were used. The value of SN for each run is printed on the output page.

The number of data points used was 1000.

x1 --	linear :	99.98
x2 --	quadratic :	94.25
x2 --	linear :	79.98
x1 --	quadratic :	66.98
linear --	quadratic :	66.74
x1 --	x2 :	27.62
x1 --	random :	25.48
quadratic --	random :	24.18
x2 --	random :	22.28
linear --	random :	19.21

Probing of K-space Relationship Discovery

Test Data Set 2

SN - 10

x1 --	quadratic :	99.96
x2 --	quadratic :	86.06
x2 --	linear :	75.30
linear --	quadratic :	73.12
x1 --	x2 :	72.17
x1 --	linear :	70.14
quadratic --	random :	69.77
linear --	random :	68.08
x2 --	random :	67.05
x1 --	random :	66.00

Probing of K-space Relationship Discovery

Test Data Set 2

SN = 0.5

x1 --	linear :	99.97
x1 --	quadratic :	90.30
x2 --	quadratic :	86.37
x2 --	linear :	86.16
linear --	quadratic :	61.44
x1 --	random :	47.50
x2 --	random :	45.19
quadratic --	random :	43.60
linear --	random :	39.19
x1 --	x2 :	38.53

Probing of K-space Relationship Discovery

Test Data Set 2

SN = 2

x1 --	quadratic :	99.96
x2 --	linear :	93.00
x2 --	quadratic :	87.78
x1 --	linear :	78.93
linear --	quadratic :	71.08
linear --	random :	62.37
x1 --	random :	60.86
x1 --	x2 :	57.88
quadratic --	random :	55.26
x2 --	random :	53.68

Probing of K-space Relationship Discovery

Test Data Set 2

SN = 1

x1 --	linear :	99.98
x2 --	quadratic :	97.02
x2 --	linear :	90.40
x1 --	quadratic :	74.94
linear --	quadratic :	70.06
x1 --	x2 :	33.85
x1 --	random :	31.39
quadratic --	random :	27.55
x2 --	random :	27.31
linear --	random :	24.79

Probing of K-space Relationship Discovery

Test Data Set 2

SN = 5

Appendix 5

Chi-squared Relationship Discovery on Civilian Customer Data

This appendix contains the output of Chi-squared Relationship Discovery run on a data set obtained from a civilian customer of HNC. This data set contained data on cellular phone usage and billing.

Also included are surface and contour plots of two relationships that were of interest to the customer.

01SEPLAN 24 --	OTACDTCL 11: 1048.11	RELATIONSHIP DETECTED
01ROMAIR --	01ROMTOL: 851.43	RELATIONSHIP DETECTED
OTAMTYPE 0 --	OTAMTYPE 2: 708.49	RELATIONSHIP DETECTED
01AMCUDU --	01AMTODU: 615.98	RELATIONSHIP DETECTED
01FTR24 --	01SEPLAN 24: 524.29	RELATIONSHIP DETECTED
01FTR24 --	OTACDTCL 11: 524.08	RELATIONSHIP DETECTED
01FTRTOT --	01RCFCHG: 427.08	RELATIONSHIP DETECTED
01FTR41 --	01FTRTOT: 389.77	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 5: 320.26	RELATIONSHIP DETECTED
01FTR99 --	01FTRTOT: 311.02	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 2: 275.69	RELATIONSHIP DETECTED
OTACDTLM --	OTNUMDAY: 266.82	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 10: 263.02	RELATIONSHIP DETECTED
01ROMAIR --	01ROMSRG: 262.79	RELATIONSHIP DETECTED
01FTR41 --	01RCFCHG: 252.60	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 16: 207.11	RELATIONSHIP DETECTED
01AMAR60 --	OTACDTCL 11: 205.28	RELATIONSHIP DETECTED
01AMAR60 --	01SEPLAN 24: 202.95	RELATIONSHIP DETECTED
01AMCUDU --	01ANPHON: 200.82	RELATIONSHIP DETECTED
01ROMSRG --	01ROMTOL: 198.05	RELATIONSHIP DETECTED
01AMCUDU --	01MINPEK: 195.58	RELATIONSHIP DETECTED
OTACDTCL 2 --	OTACDTCL 8: 195.16	RELATIONSHIP DETECTED
01AMTODU --	01ANPHON: 182.81	RELATIONSHIP DETECTED
01FTR11 --	01FTR41: 182.46	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 14: 180.53	RELATIONSHIP DETECTED
01FTR43 --	01RCFCHG: 173.67	RELATIONSHIP DETECTED
01FTR04 --	01RCFCHG: 172.74	RELATIONSHIP DETECTED
01SEPLAN 23 --	OTNUMDAY: 169.94	RELATIONSHIP DETECTED
OTAMTYPE 1 --	OTAMTYPE 2: 169.52	RELATIONSHIP DETECTED
01AMTODU --	01MINPEK: 158.10	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 11: 150.85	RELATIONSHIP DETECTED
01AMAR30 --	01AMAR60: 150.00	RELATIONSHIP DETECTED
01FTR43 --	01FTRTOT: 119.12	RELATIONSHIP DETECTED
01FTR04 --	01FTRTOT: 115.63	RELATIONSHIP DETECTED
01FTR11 --	OTACDTCL 11: 114.80	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 24: 114.16	RELATIONSHIP DETECTED
01FTR11 --	01SEPLAN 24: 113.62	RELATIONSHIP DETECTED
01ACSCHG --	OTACDTCL 11: 112.94	RELATIONSHIP DETECTED
01ACSCHG --	OTACDTCL 1: 112.30	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 6: 111.42	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 19: 106.87	RELATIONSHIP DETECTED
01FTR41 --	OTACDTCL 11: 101.40	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 24: 100.04	RELATIONSHIP DETECTED
01FTRTOT --	OTACDTCL 11: 96.43	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 24: 96.31	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 0: 96.06	RELATIONSHIP DETECTED
01AMAR60 --	01FTR24: 95.87	RELATIONSHIP DETECTED
01ACSCHG --	01AMCUDU: 93.75	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 2: 88.97	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 2: 83.72	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTACDTCL 2: 80.94	RELATIONSHIP DETECTED
01ACSCHG --	OTNUMDAY: 80.16	RELATIONSHIP DETECTED
01SEPLAN 13 --	OTACDTCL 1: 79.59	RELATIONSHIP DETECTED
01FTR11 --	01FTRTOT: 77.67	RELATIONSHIP DETECTED
01ACSCHG --	01AMTODU: 75.78	RELATIONSHIP DETECTED
01AMAR30 --	01AMTODU: 75.47	RELATIONSHIP DETECTED
01SEPLAN 14 --	OTACDTCL 1: 74.73	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTACDTCL 8: 74.19	RELATIONSHIP DETECTED
OTACDTCL 8 --	OTACDTLM: 72.11	RELATIONSHIP DETECTED
01FTR24 --	01FTRTOT: 71.52	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 23: 70.26	RELATIONSHIP DETECTED
01INTRAS --	01INTRST: 68.24	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 15: 66.30	RELATIONSHIP DETECTED
01INTRAS --	01MINPEK: 65.81	RELATIONSHIP DETECTED
01INTRST --	01MINPEK: 63.05	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTACDTLM: 62.69	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTNUMDAY: 60.94	RELATIONSHIP DETECTED
01FTR99 --	OTNUMDAY: 59.45	RELATIONSHIP DETECTED
OTACDTLM --	OTAMTYPE 2: 59.28	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 13: 58.80	RELATIONSHIP DETECTED
01ACSCHG --	01FTR24: 58.56	RELATIONSHIP DETECTED
01AMAR30 --	OTACDTCL 11: 57.55	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 20: 57.52	RELATIONSHIP DETECTED

01AMAR30	--	01SEPLAN 24:	56.69	RELATIONSHIP	DETECTED
01SEPLAN 18	--	OTNUMDAY:	56.30	RELATIONSHIP	DETECTED
01SEPLAN 2	--	OTNUMDAY:	51.97	RELATIONSHIP	DETECTED
01SEPLAN 2	--	OTACDTCL 8:	51.12	RELATIONSHIP	DETECTED
01ACSCHG	--	01SEPLAN 4:	50.90	RELATIONSHIP	DETECTED
OTACDTLM	--	OTAMTYPE 0:	50.80	RELATIONSHIP	DETECTED
01INTRAS	--	01ROMTOL:	50.78	RELATIONSHIP	DETECTED
OTACDTCL 4	--	OTACDTLM:	50.58	RELATIONSHIP	DETECTED
01AMCUDU	--	01SEPLAN 24:	49.38	RELATIONSHIP	DETECTED
OTACDTCL 2	--	OTACDTCL 4:	48.98	RELATIONSHIP	DETECTED
01INTRAS	--	01ROMAIR:	48.77	RELATIONSHIP	DETECTED
01AMAR30	--	01FTR24:	48.76	RELATIONSHIP	DETECTED
01AMAR60	--	01FTR11:	48.75	RELATIONSHIP	DETECTED
01AMCUDU	--	OTACDTCL 11:	48.51	RELATIONSHIP	DETECTED
01ACSCHG	--	01SEPLAN 3:	47.91	RELATIONSHIP	DETECTED
OTACDTCL 2	--	OTACDTLM:	46.31	RELATIONSHIP	DETECTED
01AMAR60	--	01FTR41:	46.00	RELATIONSHIP	DETECTED
OTACDTCL 2	--	OTACDTCL 13:	45.88	RELATIONSHIP	DETECTED
01SEPLAN 11	--	OTACDTCL 1:	45.81	RELATIONSHIP	DETECTED
01FTR99	--	01SEPLAN 23:	45.44	RELATIONSHIP	DETECTED
01AMCUDU	--	01INTRST:	45.00	RELATIONSHIP	DETECTED
OTACDTCL 4	--	OTACDTCL 8:	44.91	RELATIONSHIP	DETECTED
01AMCUDU	--	01INTRAS:	44.57	RELATIONSHIP	DETECTED
01SEPLAN 1	--	OTNUMDAY:	43.62	RELATIONSHIP	DETECTED
01ANPHON	--	OTACDTLM:	43.33	RELATIONSHIP	DETECTED
01MINPER	--	01SEPLAN 2:	42.67	RELATIONSHIP	DETECTED
01ANPHON	--	OTACDTCL 1:	42.64	RELATIONSHIP	DETECTED
01SEPLAN 11	--	01SEPLAN 18:	42.14	RELATIONSHIP	DETECTED
OTACDTCL 8	--	OTACDTCL 13:	42.07	RELATIONSHIP	DETECTED
01SEPLAN 17	--	OTNUMDAY:	41.74	RELATIONSHIP	DETECTED
01SEPLAN 0	--	OTNUMDAY:	41.23	RELATIONSHIP	DETECTED
01AMCUDU	--	OTACDTCL 1:	41.07	RELATIONSHIP	DETECTED
01AMTODU	--	01INTRAS:	39.54	RELATIONSHIP	DETECTED
01AMTODU	--	OTACDTCL 8:	39.47	RELATIONSHIP	DETECTED
01AMTODU	--	01INTRST:	39.23	RELATIONSHIP	DETECTED
01FTR24	--	01FTR41:	38.85	RELATIONSHIP	DETECTED
01SEPLAN 15	--	OTACDTCL 1:	38.33	RELATIONSHIP	DETECTED
OTACDTCL 8	--	OTNUMDAY:	36.23	RELATIONSHIP	DETECTED
OTACDTCL 13	--	OTNUMDAY:	36.15	RELATIONSHIP	DETECTED
OTAMTYPE 2	--	OTNUMDAY:	36.06	RELATIONSHIP	DETECTED
01ACSCHG	--	01MINPER:	35.85	RELATIONSHIP	DETECTED
01AMTODU	--	OTACDTCL 1:	34.52	RELATIONSHIP	DETECTED
01SEPLAN 0	--	OTACDTLM:	33.28	RELATIONSHIP	DETECTED
01ACSCHG	--	01SEPLAN 1:	31.89	RELATIONSHIP	DETECTED
OTAMTYPE 0	--	OTNUMDAY:	31.75	RELATIONSHIP	DETECTED
01AMTODU	--	01SEPLAN 24:	31.39	RELATIONSHIP	DETECTED
01AMTODU	--	OTACDTCL 11:	30.64	RELATIONSHIP	DETECTED
01AMCUDU	--	OTACDTCL 8:	30.58	RELATIONSHIP	DETECTED
01AMAR60	--	01FTRTOL:	29.64	RELATIONSHIP	DETECTED
01SEPLAN 5	--	OTNUMDAY:	29.60	RELATIONSHIP	DETECTED
01SEPLAN 5	--	01SEPLAN 18:	28.55	RELATIONSHIP	DETECTED
01AMCUDU	--	01FTR24:	27.66	RELATIONSHIP	DETECTED
01SEPLAN 6	--	OTNUMDAY:	26.81	RELATIONSHIP	DETECTED
01ANPHON	--	OTACDTCL 8:	26.73	RELATIONSHIP	DETECTED
OTACDTCL 13	--	OTACDTLM:	26.20	RELATIONSHIP	DETECTED
01INTRST	--	01ROMTOL:	26.16	RELATIONSHIP	DETECTED
01AMCUDU	--	OTACDTLM:	26.06	RELATIONSHIP	DETECTED
01FTR04	--	01FTR43:	25.22	RELATIONSHIP	DETECTED
01INTRST	--	01ROMAIR:	25.21	RELATIONSHIP	DETECTED
01SEPLAN 0	--	OTACDTCL 2:	25.03	RELATIONSHIP	DETECTED
01ACSCHG	--	01SEPLAN 7:	24.60	RELATIONSHIP	DETECTED
01SEPLAN 4	--	OTNUMDAY:	24.42	RELATIONSHIP	DETECTED
01SEPLAN 18	--	01SEPLAN 23:	24.28	RELATIONSHIP	DETECTED
OTACDTCL 14	--	OTNUMDAY:	24.24	RELATIONSHIP	DETECTED
01AMCUDU	--	01ROMTOL:	24.00	RELATIONSHIP	DETECTED
01FTRTOT	--	01MINPER:	23.63	RELATIONSHIP	DETECTED
01AMCUDU	--	01SEPLAN 6:	23.60	RELATIONSHIP	DETECTED
01AMTODU	--	01ROMTOL:	23.59	RELATIONSHIP	DETECTED
01SEPLAN 10	--	OTNUMDAY:	23.55	RELATIONSHIP	DETECTED
01ACSCHG	--	OTACDTCL 2:	23.18	RELATIONSHIP	DETECTED
01AMCUDU	--	01ROMAIR:	23.07	RELATIONSHIP	DETECTED
01SEPLAN 6	--	OTACDTLM:	22.48	RELATIONSHIP	DETECTED
01AMTODU	--	01ROMAIR:	22.08	RELATIONSHIP	DETECTED

01ACSCBG	--	OTACDTLM:	21.92	RELATIONSHIP	DETECTED
01AMTODU	--	01SEPLAN 6:	21.86	RELATIONSHIP	DETECTED
01AMAR30	--	OTACDTCL 8:	21.49	RELATIONSHIP	DETECTED
01MINPEK	--	OTACDTCL 8:	21.41	RELATIONSHIP	DETECTED
01SEPLAN 13	--	OTNUMDAY:	21.38	RELATIONSHIP	DETECTED
01AMTODU	--	01FTR24:	21.20	RELATIONSHIP	DETECTED
01SEPLAN 2	--	01SEPLAN 18:	21.19	RELATIONSHIP	DETECTED
01MINPEK	--	OTACDTCL 1:	20.88	RELATIONSHIP	DETECTED
01SEPLAN 5	--	01SEPLAN 11:	20.61	RELATIONSHIP	DETECTED
01AMTODU	--	OTACDTLM:	20.51	RELATIONSHIP	DETECTED
01INTRST	--	01ROMSRG:	19.72	RELATIONSHIP	DETECTED
01SEPLAN 5	--	OTACDTCL 2:	19.63	RELATIONSHIP	DETECTED
01FTRTOT	--	01SEPLAN 2:	19.45	RELATIONSHIP	DETECTED
01SEPLAN 5	--	OTACDTCL 1:	19.35	RELATIONSHIP	DETECTED
01FTR11	--	01FTR24:	19.00	RELATIONSHIP	DETECTED
OTACDTCL 1	--	OTACDTCL 4:	18.77	RELATIONSHIP	DETECTED
01ACSCBG	--	01SEPLAN 17:	18.69	RELATIONSHIP	DETECTED
OTAMTYPE 2	--	OTAMTYPE 3:	18.66	RELATIONSHIP	DETECTED
01FTRTOT	--	OTNUMDAY:	18.28	RELATIONSHIP	DETECTED
01AMTODU	--	01RCFCHG:	18.14	RELATIONSHIP	DETECTED
01AMCUDU	--	01RCFCHG:	17.77	RELATIONSHIP	DETECTED
OTACDTCL 1	--	OTACDTCL 13:	17.59	RELATIONSHIP	DETECTED
01SEPLAN 11	--	01SEPLAN 23:	17.54	RELATIONSHIP	DETECTED
01SEPLAN 11	--	OTNUMDAY:	17.48	RELATIONSHIP	DETECTED
OTACDTCL 2	--	OTNUMDAY:	17.28	RELATIONSHIP	DETECTED
OTACDTCL 2	--	OTACDTCL 11:	16.74	RELATIONSHIP	DETECTED
01ACSCBG	--	OTACDTCL 13:	16.73	RELATIONSHIP	DETECTED
01AMAR30	--	01FTR41:	16.71	RELATIONSHIP	DETECTED
01FTR11	--	01RCFCHG:	16.62	RELATIONSHIP	DETECTED
01FTR10	--	01FTRTOT:	16.46	RELATIONSHIP	DETECTED
01SEPLAN 24	--	OTACDTCL 2:	16.37	RELATIONSHIP	DETECTED
01MINPEK	--	01RCFCHG:	16.15	RELATIONSHIP	DETECTED
01ACSCBG	--	01AMAR60:	15.79	RELATIONSHIP	DETECTED
01AMAR60	--	01AMTODU:	15.77	RELATIONSHIP	DETECTED
01MINPEK	--	01SEPLAN 10:	15.74	RELATIONSHIP	DETECTED
01SEPLAN 24	--	OTACDTCL 8:	15.53	RELATIONSHIP	DETECTED
OTACDTCL 8	--	OTACDTCL 11:	15.37	RELATIONSHIP	DETECTED
01SEPLAN 2	--	01SEPLAN 11:	15.32	RELATIONSHIP	DETECTED
01AMAR30	--	OTACDTCL 4:	14.92	RELATIONSHIP	DETECTED
01SEPLAN 23	--	OTACDTLM:	14.91	RELATIONSHIP	DETECTED
01RCFCHG	--	01SEPLAN 18:	14.75	RELATIONSHIP	DETECTED
01SEPLAN 2	--	OTACDTCL 1:	14.74	RELATIONSHIP	DETECTED
01SEPLAN 13	--	OTACDTLM:	14.30	RELATIONSHIP	DETECTED
01FTR13	--	01RCFCHG:	14.14	RELATIONSHIP	DETECTED
01MINPEK	--	01SEPLAN 6:	14.05	RELATIONSHIP	DETECTED
01SEPLAN 14	--	01SEPLAN 18:	13.88	RELATIONSHIP	DETECTED
01FTR11	--	OTNUMDAY:	13.72	RELATIONSHIP	DETECTED
01MINPEK	--	01ROMTOL:	13.30	RELATIONSHIP	DETECTED
01MINPEK	--	01ROMAIR:	13.25	RELATIONSHIP	DETECTED
01ACSCBG	--	01FTR11:	13.21	RELATIONSHIP	DETECTED
01ACSCBG	--	01SEPLAN 22:	13.20	RELATIONSHIP	DETECTED
01AMAR30	--	01FTRTOT:	13.12	RELATIONSHIP	DETECTED
01FTRTOT	--	OTACDTCL 8:	12.99	RELATIONSHIP	DETECTED
01SEPLAN 23	--	OTACDTCL 13:	12.86	RELATIONSHIP	DETECTED
01ACSCBG	--	01FTR99:	12.79	RELATIONSHIP	DETECTED
01FTR99	--	01SEPLAN 25:	12.61	RELATIONSHIP	DETECTED
01AMSTAT 0	--	01AMSTAT 1:	12.59	RELATIONSHIP	DETECTED
01SEPLAN 12	--	OTACDTCL 1:	12.53	RELATIONSHIP	DETECTED
01SEPLAN 6	--	OTAMTYPE 0:	12.52	RELATIONSHIP	DETECTED
01FTR41	--	01FTR85:	12.36	RELATIONSHIP	DETECTED
01INTRAS	--	01SEPLAN 2:	12.29	RELATIONSHIP	DETECTED
01AMAR60	--	OTACDTCL 8:	12.04	RELATIONSHIP	DETECTED
01SEPLAN 5	--	01SEPLAN 23:	11.94	RELATIONSHIP	DETECTED
01SEPLAN 10	--	01SEPLAN 18:	11.93	RELATIONSHIP	DETECTED
01AMCUDU	--	01SEPLAN 18:	11.81	RELATIONSHIP	DETECTED
01AMTODU	--	01FTRTOT:	11.80	RELATIONSHIP	DETECTED
01SEPLAN 6	--	OTACDTCL 2:	11.74	RELATIONSHIP	DETECTED
01AMAR30	--	OTACDTCL 13:	11.74	RELATIONSHIP	DETECTED
01FTR41	--	01MINPEK:	11.72	RELATIONSHIP	DETECTED
01SEPLAN 14	--	OTACDTLM:	11.64	RELATIONSHIP	DETECTED
01ACSCBG	--	01INTRAS:	11.60	RELATIONSHIP	DETECTED
01FTR03	--	01RCFCHG:	11.54	RELATIONSHIP	DETECTED
01SEPLAN 21	--	OTNUMDAY:	11.46	RELATIONSHIP	DETECTED

01AMCUDU --	01ROMSRG: 11.46	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 8: 11.46	RELATIONSHIP DETECTED
01ACSCHG --	01RCFCHG: 11.44	RELATIONSHIP DETECTED
01SEPLAN 6 --	OTAMTYPE 2: 11.35	RELATIONSHIP DETECTED
01SEPLAN 22 --	OTNUMDAY: 11.04	RELATIONSHIP DETECTED
01AMAR30 --	01AMCUDU: 11.01	RELATIONSHIP DETECTED
01LCLLND --	01MINPEK: 10.98	RELATIONSHIP DETECTED
01ACSCHG --	01ROMAIR: 10.89	RELATIONSHIP DETECTED
01SEPLAN 23 --	OTACDTCL 1: 10.88	RELATIONSHIP DETECTED
01FTR13 --	01FTRTOT: 10.88	RELATIONSHIP DETECTED
01FTR10 --	01RCFCHG: 10.78	RELATIONSHIP DETECTED
OTACDTCL 4 --	OTACDTCL 13: 10.75	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 18: 10.50	RELATIONSHIP DETECTED
OTACDTLM --	OTAMTYPE 1: 10.46	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 5: 10.44	RELATIONSHIP DETECTED
01SEPLAN 14 --	OTNUMDAY: 10.38	RELATIONSHIP DETECTED
01AMCUDU --	01FTRTOT: 10.11	RELATIONSHIP DETECTED
01ACSCHG --	01ANPHON: 10.08	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 14: 10.07	RELATIONSHIP DETECTED
01SEPLAN 25 --	OTACDTCL 6: 10.03	RELATIONSHIP DETECTED
01SEPLAN 10 --	OTACDTCL 13: 9.92	RELATIONSHIP DETECTED
01INTRST --	01SEPLAN 2: 9.81	RELATIONSHIP DETECTED
01AMTODU --	01ROMSRG: 9.77	RELATIONSHIP DETECTED
01SEPLAN 25 --	OTACDTCL 1: 9.76	RELATIONSHIP DETECTED
01FTR41 --	OTACDTCL 8: 9.70	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 25: 9.61	RELATIONSHIP DETECTED
01ACSCHG --	01INTRST: 9.55	RELATIONSHIP DETECTED
01AMCUDU --	OTNUMDAY: 9.38	RELATIONSHIP DETECTED
01LCLLND --	01ROMAIR: 9.31	RELATIONSHIP DETECTED
01SEPLAN 18 --	OTACDTLM: 9.28	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 25: 9.15	RELATIONSHIP DETECTED
01ACSCHG --	OTACDTCL 8: 9.13	RELATIONSHIP DETECTED
01FTR85 --	01SEPLAN 17: 9.08	RELATIONSHIP DETECTED
01AMAR30 --	01FTR11: 9.07	RELATIONSHIP DETECTED
01FTR24 --	OTACDTCL 2: 9.05	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 2: 9.02	RELATIONSHIP DETECTED
01SEPLAN 18 --	01SEPLAN 24: 9.00	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 17: 8.96	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 18: 8.94	RELATIONSHIP DETECTED
01SEPLAN 18 --	OTACDTCL 11: 8.91	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 23: 8.91	RELATIONSHIP DETECTED
01FTR19 --	01FTR41: 8.90	RELATIONSHIP DETECTED
01MINPEK --	01SEPLAN 18: 8.86	RELATIONSHIP DETECTED
01SEPLAN 6 --	01SEPLAN 18: 8.80	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 9: 8.78	RELATIONSHIP DETECTED
01LCLLND --	01ROMTOL: 8.72	RELATIONSHIP DETECTED
01FTR04 --	01MINPEK: 8.68	RELATIONSHIP DETECTED
01SEPLAN 10 --	01SEPLAN 11: 8.66	RELATIONSHIP DETECTED
01FTR03 --	01FTRTOT: 8.61	RELATIONSHIP DETECTED
01SEPLAN 10 --	OTACDTCL 1: 8.51	RELATIONSHIP DETECTED
01FTR19 --	01SEPLAN 25: 8.51	RELATIONSHIP DETECTED
01ACSCHG --	01ROMTOL: 8.45	RELATIONSHIP DETECTED
01ACSCHG --	OTACDTCL 4: 8.38	RELATIONSHIP DETECTED
01SEPLAN 18 --	01SEPLAN 19: 8.36	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 12: 8.31	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 11: 8.28	RELATIONSHIP DETECTED
01RCFCHG --	01SEPLAN 2: 8.27	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 10: 8.14	RELATIONSHIP DETECTED
01FTR24 --	OTACDTCL 8: 8.05	RELATIONSHIP DETECTED
01MINPEK --	OTNUMDAY: 7.91	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 21: 7.89	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 10: 7.88	RELATIONSHIP DETECTED
01FTR99 --	01FTR41: 7.84	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 22: 7.78	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 23: 7.72	RELATIONSHIP DETECTED
01FTR85 --	01FTR41: 7.71	RELATIONSHIP DETECTED
01RCFCHG --	01FTRTOT: 7.64	RELATIONSHIP DETECTED
01ACSCHG --	OTNUMDAY: 7.60	RELATIONSHIP DETECTED
01MINPEK --	01FTRTOT: 7.58	RELATIONSHIP DETECTED
01FTR99 --	OTACDTLM: 7.57	RELATIONSHIP DETECTED
01ACSCHG --	01SEPLAN 1: 7.49	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 16: 7.47	RELATIONSHIP DETECTED
	OTACDTLM: 7.30	RELATIONSHIP DETECTED

01AMCUDU --	01LCLLND: 7.23	RELATIONSHIP DETECTED
01FTR11 --	01MINPEK: 7.21	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 25: 7.18	RELATIONSHIP DETECTED
01FTR11 --	OTACDTCL 8: 7.05	RELATIONSHIP DETECTED
01SEPLAN 13 --	OTACDTCL 2: 7.02	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 2: 6.99	RELATIONSHIP DETECTED
01SEPLAN 17 --	01SEPLAN 18: 6.98	RELATIONSHIP DETECTED
01SEPLAN 3 --	OTNUMDAY: 6.93	RELATIONSHIP DETECTED
OTACDTCL 2 --	OTACDTCL 12: 6.91	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 14: 6.90	RELATIONSHIP DETECTED
01ACSCHG --	01AMAR30: 6.85	RELATIONSHIP DETECTED
01AMTODU --	OTNUMDAY: 6.84	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 5: 6.83	RELATIONSHIP DETECTED
01SEPLAN 7 --	OTNUMDAY: 6.81	RELATIONSHIP DETECTED
01MINPEK --	01SEPLAN 7: 6.79	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 5: 6.76	RELATIONSHIP DETECTED
OTACDTCL 2 --	OTACDTCL 3: 6.70	RELATIONSHIP DETECTED
01SEPLAN 18 --	OTACDTCL 1: 6.66	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 25: 6.64	RELATIONSHIP DETECTED
01SEPLAN 24 --	OTACDTCL 1: 6.61	RELATIONSHIP DETECTED
01SEPLAN 13 --	OTACDTCL 8: 6.60	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 24: 6.56	RELATIONSHIP DETECTED
01SEPLAN 0 --	OTACDTCL 1: 6.56	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTACDTCL 11: 6.55	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 25: 6.54	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 11: 6.51	RELATIONSHIP DETECTED
01SEPLAN 11 --	OTACDTCL 11: 6.50	RELATIONSHIP DETECTED
01INTRST --	01RCFCHG: 6.50	RELATIONSHIP DETECTED
01FTR10 --	01SEPLAN 24: 6.49	RELATIONSHIP DETECTED
OTACDTCL 2 --	OTACDTCL 14: 6.44	RELATIONSHIP DETECTED
01FTR11 --	OTACDTCL 6: 6.44	RELATIONSHIP DETECTED
01SEPLAN 11 --	OTACDTCL 2: 6.42	RELATIONSHIP DETECTED
01FTR11 --	01SEPLAN 25: 6.42	RELATIONSHIP DETECTED
01SEPLAN 6 --	01SEPLAN 11: 6.41	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 6: 6.38	RELATIONSHIP DETECTED
OTACDTCL 8 --	OTACDTCL 12: 6.36	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 9: 6.34	RELATIONSHIP DETECTED
01FTR41 --	OTNUMDAY: 6.34	RELATIONSHIP DETECTED
01RCFCHG --	OTACDTIM: 6.33	RELATIONSHIP DETECTED
01AMAR30 --	01SEPLAN 2: 6.33	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 7: 6.31	RELATIONSHIP DETECTED
01SEPLAN 18 --	OTACDTCL 13: 6.30	RELATIONSHIP DETECTED
01AMAR60 --	OTACDTCL 2: 6.26	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 7: 6.25	RELATIONSHIP DETECTED
01SEPLAN 6 --	OTACDTCL 1: 6.24	RELATIONSHIP DETECTED
01AMAR30 --	01MINPEK: 6.23	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 5: 6.19	RELATIONSHIP DETECTED
OTACDTCL 3 --	OTACDTCL 8: 6.16	RELATIONSHIP DETECTED
01AMAR30 --	01RCFCHG: 6.16	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 19: 6.10	RELATIONSHIP DETECTED
01AMCUDU --	01FTR41: 6.10	RELATIONSHIP DETECTED
01SEPLAN 15 --	OTACDTIM: 6.06	RELATIONSHIP DETECTED
01SEPLAN 10 --	OTACDTIM: 6.06	RELATIONSHIP DETECTED
01FTR43 --	01MINPEK: 6.01	RELATIONSHIP DETECTED
01AMTODU --	01LCLLND: 5.98	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 10: 5.95	RELATIONSHIP DETECTED
OTACDTCL 8 --	OTACDTCL 14: 5.93	RELATIONSHIP DETECTED
01ANPHON --	01FTR24: 5.93	RELATIONSHIP DETECTED
01RCFCHG --	OTACDTCL 8: 5.92	RELATIONSHIP DETECTED
01ANPHON --	OTACDTCL 11: 5.92	RELATIONSHIP DETECTED
01SEPLAN 14 --	01SEPLAN 23: 5.91	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 12: 5.91	RELATIONSHIP DETECTED
OTACDTCL 2 --	OTACDTCL 5: 5.90	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 11: 5.90	RELATIONSHIP DETECTED
01FTR42 --	01RCFCHG: 5.89	RELATIONSHIP DETECTED
01ROMAIR --	01SEPLAN 2: 5.83	RELATIONSHIP DETECTED
01SEPLAN 14 --	OTACDTCL 2: 5.79	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 25: 5.78	RELATIONSHIP DETECTED
01AMTODU --	01FTR04: 5.75	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTAMTYPE 2: 5.72	RELATIONSHIP DETECTED
01AMTODU --	OTACDTCL 4: 5.72	RELATIONSHIP DETECTED
01AMCUDU --	01FTR99: 5.71	RELATIONSHIP DETECTED
01FTR19 --	01FTRTOT: 5.71	RELATIONSHIP DETECTED

OTACDTCL 2 --	OTACDTCL 10:	5.69	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 24:	5.69	RELATIONSHIP DETECTED
01INTRAS --	01RCFCHG:	5.66	RELATIONSHIP DETECTED
01SEPLAN 3 --	01SEPLAN 18:	5.65	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 23:	5.64	RELATIONSHIP DETECTED
01FTR10 --	OTACDTCL 11:	5.63	RELATIONSHIP DETECTED
01FTR41 --	OTACDTCL 6:	5.60	RELATIONSHIP DETECTED
01ANPHON --	OTACDTCL 10:	5.56	RELATIONSHIP DETECTED
01FTR04 --	01INTRST:	5.56	RELATIONSHIP DETECTED
01INTRST --	01LCLLND:	5.54	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 5:	5.44	RELATIONSHIP DETECTED
OTACDTCL 5 --	OTACDTCL 8:	5.43	RELATIONSHIP DETECTED
01AMAR30 --	01SEPLAN 18:	5.43	RELATIONSHIP DETECTED
01SEPLAN 15 --	01SEPLAN 18:	5.41	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 18:	5.34	RELATIONSHIP DETECTED
01MINPER --	01SEPLAN 11:	5.34	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 23:	5.34	RELATIONSHIP DETECTED
01ROMTOL --	01SEPLAN 2:	5.33	RELATIONSHIP DETECTED
OTACDTCL 8 --	OTACDTCL 10:	5.24	RELATIONSHIP DETECTED
OTAMTYPE 1 --	OTNUMDAY:	5.18	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 14:	5.18	RELATIONSHIP DETECTED
01FTRTOT --	01INTRAS:	5.16	RELATIONSHIP DETECTED
01ANPHON --	OTACDTCL 13:	5.16	RELATIONSHIP DETECTED
01AMCUDU --	01FTR04:	5.14	RELATIONSHIP DETECTED
01INTRAS --	OTACDTCL 8:	5.13	RELATIONSHIP DETECTED
01AMCUDU --	01FTR11:	5.13	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 17:	5.11	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 13:	5.11	RELATIONSHIP DETECTED
01SEPLAN 10 --	01SEPLAN 23:	5.10	RELATIONSHIP DETECTED
01FTR24 --	OTACDTIM:	5.10	RELATIONSHIP DETECTED
01FTRTOT --	01INTRST:	5.09	RELATIONSHIP DETECTED
01FTR24 --	01SEPLAN 18:	5.03	RELATIONSHIP DETECTED
OTACDTCL 12 --	OTNUMDAY:	4.98	RELATIONSHIP DETECTED
01INTRAS --	OTACDTCL 1:	4.97	RELATIONSHIP DETECTED
01AMTODU --	01FTR99:	4.96	RELATIONSHIP DETECTED
01FTR43 --	OTACDTIM:	4.89	RELATIONSHIP DETECTED
01INTRAS --	01ROMSRG:	4.86	RELATIONSHIP DETECTED
01RCFCHG --	01SEPLAN 11:	4.83	RELATIONSHIP DETECTED
01INTRAS --	01SEPLAN 6:	4.82	RELATIONSHIP DETECTED
OTACDTCL 8 --	OTAMTYPE 2:	4.79	RELATIONSHIP DETECTED
01SEPLAN 13 --	01SEPLAN 18:	4.76	RELATIONSHIP DETECTED
01SEPLAN 4 --	01SEPLAN 18:	4.76	RELATIONSHIP DETECTED
OTACDTCL 3 --	OTACDTIM:	4.74	RELATIONSHIP DETECTED
01SEPLAN 18 --	01SEPLAN 20:	4.72	RELATIONSHIP DETECTED
01SEPLAN 2 --	OTACDTCL 13:	4.72	RELATIONSHIP DETECTED
OTACDTCL 10 --	OTACDTIM:	4.69	RELATIONSHIP DETECTED
01LCLLND --	01ROMSRG:	4.65	RELATIONSHIP DETECTED
01FTR03 --	01FTR04:	4.65	RELATIONSHIP DETECTED
01FTR04 --	OTACDTCL 8:	4.59	RELATIONSHIP DETECTED
01FTR43 --	OTNUMDAY:	4.56	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 24:	4.53	RELATIONSHIP DETECTED
01FTR99 --	OTACDTIM:	4.52	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 5:	4.50	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 23:	4.50	RELATIONSHIP DETECTED
01SEPLAN 5 --	OTACDTCL 11:	4.49	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 10:	4.48	RELATIONSHIP DETECTED
01FTR42 --	01FTRTOT:	4.45	RELATIONSHIP DETECTED
01SEPLAN 18 --	01SEPLAN 25:	4.43	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 6:	4.43	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 5:	4.43	RELATIONSHIP DETECTED
01MINPER --	01SEPLAN 24:	4.42	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 2:	4.40	RELATIONSHIP DETECTED
01FTR04 --	01INTRAS:	4.39	RELATIONSHIP DETECTED
01AMCUDU --	OTACDTCL 13:	4.34	RELATIONSHIP DETECTED
01MINPER --	OTACDTCL 11:	4.33	RELATIONSHIP DETECTED
OTACDTCL 2 --	OTAMTYPE 0:	4.28	RELATIONSHIP DETECTED
01MINPER --	01SEPLAN 13:	4.23	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 19:	4.22	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 10:	4.22	RELATIONSHIP DETECTED
01FTR24 --	01INTRAS:	4.19	RELATIONSHIP DETECTED
01INTRST --	01SEPLAN 6:	4.18	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 9:	4.18	RELATIONSHIP DETECTED
OTACDTCL 4 --	OTNUMDAY:	4.17	RELATIONSHIP DETECTED

01SEPLAN 3 --	01SEPLAN 11:	4.16	RELATIONSHIP DETECTED
01FTR11 --	OTACDTLM:	4.14	RELATIONSHIP DETECTED
01SEPLAN 20 --	OTNUMDAY:	4.13	RELATIONSHIP DETECTED
01SEPLAN 24 --	OTACDTCL 4:	4.12	RELATIONSHIP DETECTED
01SEPLAN 16 --	OTNUMDAY:	4.11	RELATIONSHIP DETECTED
OTACDTCL 4 --	OTACDTCL 11:	4.08	RELATIONSHIP DETECTED
01FTR20 --	01RCFCHG:	4.05	RELATIONSHIP DETECTED
01FTR19 --	01FTR99:	4.05	RELATIONSHIP DETECTED
OTACDTCL 5 --	OTNUMDAY:	4.01	RELATIONSHIP DETECTED
OTACDTCL 12 --	OTACDTLM:	4.00	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 15:	3.98	RELATIONSHIP DETECTED
01SEPLAN 1 --	OTACDTCL 1:	3.96	RELATIONSHIP DETECTED
01SEPLAN 15 --	OTNUMDAY:	3.94	RELATIONSHIP DETECTED
01SEPLAN 0 --	OTAMTYPE 2:	3.94	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 11:	3.93	RELATIONSHIP DETECTED
01SEPLAN 7 --	OTACDTCL 8:	3.92	RELATIONSHIP DETECTED
01ROMAIR --	01SEPLAN 18:	3.92	RELATIONSHIP DETECTED
01MINPER --	01SEPLAN 14:	3.92	RELATIONSHIP DETECTED
OTACDTCL 14 --	OTACDTLM:	3.91	RELATIONSHIP DETECTED
01SEPLAN 23 --	01SEPLAN 24:	3.90	RELATIONSHIP DETECTED
01SEPLAN 24 --	OTACDTCL 13:	3.88	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 23:	3.87	RELATIONSHIP DETECTED
01SEPLAN 23 --	OTACDTCL 11:	3.86	RELATIONSHIP DETECTED
OTACDTCL 11 --	OTACDTCL 13:	3.84	RELATIONSHIP DETECTED
01FTR24 --	01INTRST:	3.82	RELATIONSHIP DETECTED
01SEPLAN 6 --	01SEPLAN 23:	3.81	RELATIONSHIP DETECTED
01AMTODU --	01FTR11:	3.81	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 18:	3.81	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 3:	3.80	RELATIONSHIP DETECTED
01SEPLAN 0 --	OTAMTYPE 0:	3.79	RELATIONSHIP DETECTED
01SEPLAN 19 --	OTNUMDAY:	3.78	RELATIONSHIP DETECTED
01AMCUDU --	OTACDTCL 2:	3.78	RELATIONSHIP DETECTED
OTACDTCL 8 --	OTAMTYPE 0:	3.76	RELATIONSHIP DETECTED
01ROMAIR --	01SEPLAN 6:	3.76	RELATIONSHIP DETECTED
01ANPHON --	01FTRTOT:	3.76	RELATIONSHIP DETECTED
01AMTODU --	OTACDTCL 13:	3.75	RELATIONSHIP DETECTED
01FTR41 --	01FTR43:	3.74	RELATIONSHIP DETECTED
01FTR24 --	OTACDTCL 1:	3.73	RELATIONSHIP DETECTED
OTACDTCL 2 --	OTAMTYPE 2:	3.71	RELATIONSHIP DETECTED
01AMAR60 --	01ANPHON:	3.70	RELATIONSHIP DETECTED
OTACDTCL 3 --	OTNUMDAY:	3.69	RELATIONSHIP DETECTED
01ROMTOL --	01SEPLAN 18:	3.66	RELATIONSHIP DETECTED
01SEPLAN 14 --	OTACDTCL 8:	3.65	RELATIONSHIP DETECTED
01SEPLAN 0 --	OTACDTCL 13:	3.65	RELATIONSHIP DETECTED
01MINPER --	01SEPLAN 9:	3.65	RELATIONSHIP DETECTED
01SEPLAN 10 --	OTACDTCL 5:	3.64	RELATIONSHIP DETECTED
01FTR04 --	01FTR41:	3.64	RELATIONSHIP DETECTED
01SEPLAN 19 --	01SEPLAN 23:	3.63	RELATIONSHIP DETECTED
01SEPLAN 15 --	OTACDTCL 2:	3.62	RELATIONSHIP DETECTED
01AMAR60 --	01AMAR90:	3.61	RELATIONSHIP DETECTED
01FTR12 --	01RCFCHG:	3.61	RELATIONSHIP DETECTED
OTACDTCL 5 --	OTACDTLM:	3.58	RELATIONSHIP DETECTED
01SEPLAN 3 --	OTACDTCL 1:	3.56	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 17:	3.55	RELATIONSHIP DETECTED
01SEPLAN 4 --	OTACDTCL 1:	3.54	RELATIONSHIP DETECTED
01AMCUDU --	OTACDTCL 4:	3.53	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 13:	3.52	RELATIONSHIP DETECTED
01SEPLAN 4 --	01SEPLAN 11:	3.52	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 8:	3.52	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 20:	3.48	RELATIONSHIP DETECTED
01FTRTOT --	OTACDTCL 2:	3.45	RELATIONSHIP DETECTED
01ACSCHG --	01FTR19:	3.45	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 24:	3.44	RELATIONSHIP DETECTED
01INTRST --	OTACDTCL 1:	3.44	RELATIONSHIP DETECTED
01AMAR60 --	01SEPLAN 2:	3.44	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 2:	3.41	RELATIONSHIP DETECTED
01SEPLAN 2 --	OTACDTCL 11:	3.40	RELATIONSHIP DETECTED
01SEPLAN 18 --	OTACDTCL 4:	3.38	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTAMTYPE 1:	3.37	RELATIONSHIP DETECTED
01AMTODU --	OTACDTCL 10:	3.37	RELATIONSHIP DETECTED
01AMTODU --	OTACDTCL 2:	3.37	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 6:	3.36	RELATIONSHIP DETECTED
01AMAR60 --	01MINPER:	3.35	RELATIONSHIP DETECTED

01SEPLAN 13 --	OTACDTCL 0:	3.32	RELATIONSHIP DETECTED
01AMCUDU --	OTACDTCL 10:	3.32	RELATIONSHIP DETECTED
01INTRST --	OTACDTCL 8:	3.31	RELATIONSHIP DETECTED
01FTR24 --	01SEPLAN 11:	3.29	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 25:	3.28	RELATIONSHIP DETECTED
01AMAR60 --	01AMCUDU:	3.26	RELATIONSHIP DETECTED
01SEPLAN 2 --	OTACDTCL 4:	3.24	RELATIONSHIP DETECTED
01SEPLAN 0 --	OTACDTCL 8:	3.23	RELATIONSHIP DETECTED
01FTR47 --	01FTRTOT:	3.22	RELATIONSHIP DETECTED
01FTR20 --	01FTRTOT:	3.22	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 19:	3.21	RELATIONSHIP DETECTED
01AMAR60 --	OTACDTCL 13:	3.19	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 1:	3.17	RELATIONSHIP DETECTED
01SEPLAN 19 --	OTACDTCL 10:	3.16	RELATIONSHIP DETECTED
01SEPLAN 14 --	OTACDTCL 13:	3.16	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 24:	3.16	RELATIONSHIP DETECTED
01FTR99 --	OTACDTCL 11:	3.13	RELATIONSHIP DETECTED
01FTR85 --	01SEPLAN 16:	3.13	RELATIONSHIP DETECTED
01RCFCHG --	01SEPLAN 24:	3.10	RELATIONSHIP DETECTED
01ACSCHG --	01LCLLND:	3.10	RELATIONSHIP DETECTED
01SEPLAN 7 --	01SEPLAN 18:	3.09	RELATIONSHIP DETECTED
01ACSCHG --	OTAMTYPE 0:	3.09	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 6:	3.09	RELATIONSHIP DETECTED
01SEPLAN 17 --	01SEPLAN 23:	3.06	RELATIONSHIP DETECTED
01RCFCHG --	OTACDTCL 11:	3.06	RELATIONSHIP DETECTED
01SEPLAN 10 --	01SEPLAN 14:	3.03	RELATIONSHIP DETECTED
01FTR04 --	OTNUMDAY:	3.02	RELATIONSHIP DETECTED
OTACDTCL 10 --	OTNUMDAY:	3.01	RELATIONSHIP DETECTED
01ACSCHG --	01FTR04:	3.01	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 9:	3.00	RELATIONSHIP DETECTED
01AMAR30 --	OTNUMDAY:	3.00	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 1:	2.99	RELATIONSHIP DETECTED
01FTR99 --	OTACDTCL 13:	2.97	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTAMTYPE 0:	2.96	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 5:	2.92	RELATIONSHIP DETECTED
01SEPLAN 3 --	01SEPLAN 5:	2.91	RELATIONSHIP DETECTED
01FTRTOT --	OTACDTCL 6:	2.91	RELATIONSHIP DETECTED
01FTR12 --	01FTRTOT:	2.91	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 15:	2.89	RELATIONSHIP DETECTED
01AMAR30 --	01ROMTOL:	2.88	RELATIONSHIP DETECTED
01SEPLAN 20 --	OTACDTCL 1:	2.87	RELATIONSHIP DETECTED
01AMAR30 --	OTACDTLM:	2.86	RELATIONSHIP DETECTED
01ACSCHG --	OTAMTYPE 2:	2.84	RELATIONSHIP DETECTED
01AMAR60 --	OTNUMDAY:	2.83	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTACDTCL 12:	2.82	RELATIONSHIP DETECTED
01FTR85 --	OTNUMDAY:	2.80	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 15:	2.79	RELATIONSHIP DETECTED
01SEPLAN 3 --	OTACDTLM:	2.79	RELATIONSHIP DETECTED
01ACSCHG --	01ROMSRG:	2.79	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 5:	2.76	RELATIONSHIP DETECTED
01FTR01 --	01RCFCHG:	2.76	RELATIONSHIP DETECTED
01ROMTOL --	01SEPLAN 6:	2.75	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTACDTCL 3:	2.74	RELATIONSHIP DETECTED
01RCFCHG --	OTACDTCL 1:	2.74	RELATIONSHIP DETECTED
01AMAR60 --	01SEPLAN 23:	2.74	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 17:	2.71	RELATIONSHIP DETECTED
01SEPLAN 1 --	OTACDTLM:	2.71	RELATIONSHIP DETECTED
01FTR20 --	01FTR41:	2.71	RELATIONSHIP DETECTED
01RCFCHG --	OTACDTCL 13:	2.70	RELATIONSHIP DETECTED
01FTR04 --	OTACDTLM:	2.70	RELATIONSHIP DETECTED
OTACDTCL 13 --	OTAMTYPE 2:	2.69	RELATIONSHIP DETECTED
01AMAR30 --	01ANPHON:	2.69	RELATIONSHIP DETECTED
01SEPLAN 24 --	OTNUMDAY:	2.68	RELATIONSHIP DETECTED
01SEPLAN 12 --	01SEPLAN 18:	2.68	RELATIONSHIP DETECTED
01SEPLAN 6 --	OTACDTCL 13:	2.68	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTACDTCL 14:	2.65	RELATIONSHIP DETECTED
01ANPHON --	01MINPEK:	2.65	RELATIONSHIP DETECTED
01FTR19 --	OTNUMDAY:	2.65	RELATIONSHIP DETECTED
01MINPEK --	01SEPLAN 8:	2.64	RELATIONSHIP DETECTED
OTACDTCL 11 --	OTNUMDAY:	2.63	RELATIONSHIP DETECTED
01AMCUDU --	01FTR43:	2.62	RELATIONSHIP DETECTED
01FTR24 --	01SEPLAN 5:	2.61	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 14:	2.60	RELATIONSHIP DETECTED

01MINPER --	01SEPLAN 5: 2.58	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 0: 2.58	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 6: 2.57	RELATIONSHIP DETECTED
01FTR41 --	01INTRST: 2.57	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 0: 2.56	RELATIONSHIP DETECTED
01SEPLAN 25 --	OTNUMDAY: 2.55	RELATIONSHIP DETECTED
01FTR41 --	OTACDTCL 2: 2.55	RELATIONSHIP DETECTED
01ROMTOL --	01SEPLAN 4: 2.54	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 14: 2.53	RELATIONSHIP DETECTED
01SEPLAN 10 --	OTACDTCL 4: 2.52	RELATIONSHIP DETECTED
01SEPLAN 3 --	01SEPLAN 23: 2.52	RELATIONSHIP DETECTED
01SEPLAN 7 --	OTACDTLM: 2.51	RELATIONSHIP DETECTED
01LCLLND --	01SEPLAN 18: 2.50	RELATIONSHIP DETECTED
01INTRAS --	01SEPLAN 18: 2.50	RELATIONSHIP DETECTED
01ANPHON --	OTACDTCL 5: 2.49	RELATIONSHIP DETECTED
01FTR20 --	OTACDTCL 11: 2.49	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 13: 2.48	RELATIONSHIP DETECTED
01SEPLAN 4 --	01SEPLAN 5: 2.48	RELATIONSHIP DETECTED
01FTR20 --	01SEPLAN 24: 2.48	RELATIONSHIP DETECTED
01FTR04 --	01SEPLAN 2: 2.48	RELATIONSHIP DETECTED
01FTR99 --	01RCFCHG: 2.47	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 20: 2.46	RELATIONSHIP DETECTED
01SEPLAN 2 --	OTACDTLM: 2.45	RELATIONSHIP DETECTED
01INTRST --	OTACDTCL 4: 2.45	RELATIONSHIP DETECTED
01SEPLAN 24 --	OTACDTLM: 2.44	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTACDTCL 5: 2.44	RELATIONSHIP DETECTED
01SEPLAN 15 --	01SEPLAN 23: 2.42	RELATIONSHIP DETECTED
01LCLLND --	01SEPLAN 2: 2.42	RELATIONSHIP DETECTED
01FTR04 --	OTACDTCL 13: 2.41	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 23: 2.39	RELATIONSHIP DETECTED
01ROMAIR --	01SEPLAN 5: 2.38	RELATIONSHIP DETECTED
01ANPHON --	OTACDTCL 12: 2.38	RELATIONSHIP DETECTED
01FTR24 --	OTACDTCL 4: 2.38	RELATIONSHIP DETECTED
OTACDTCL 11 --	OTACDTLM: 2.36	RELATIONSHIP DETECTED
OTACDTCL 1 --	OTACDTCL 10: 2.36	RELATIONSHIP DETECTED
01ROMSRG --	01SEPLAN 2: 2.36	RELATIONSHIP DETECTED
01FTR24 --	01MINPER: 2.36	RELATIONSHIP DETECTED
01SEPLAN 25 --	OTACDTLM: 2.35	RELATIONSHIP DETECTED
01SEPLAN 14 --	01SEPLAN 24: 2.35	RELATIONSHIP DETECTED
01RCFCHG --	01ROMTOL: 2.34	RELATIONSHIP DETECTED
01SEPLAN 14 --	OTACDTCL 11: 2.33	RELATIONSHIP DETECTED
01SEPLAN 7 --	OTACDTCL 1: 2.33	RELATIONSHIP DETECTED
01SEPLAN 5 --	OTACDTCL 13: 2.33	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 14: 2.33	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 10: 2.33	RELATIONSHIP DETECTED
01SEPLAN 23 --	OTACDTCL 12: 2.32	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 25: 2.32	RELATIONSHIP DETECTED
01FTR01 --	01FTRTOT: 2.32	RELATIONSHIP DETECTED
01SEPLAN 7 --	01SEPLAN 11: 2.31	RELATIONSHIP DETECTED
01SEPLAN 6 --	01SEPLAN 14: 2.30	RELATIONSHIP DETECTED
01SEPLAN 15 --	OTACDTCL 8: 2.29	RELATIONSHIP DETECTED
01FTR13 --	01FTR99: 2.29	RELATIONSHIP DETECTED
01SEPLAN 13 --	OTAMTYPE 0: 2.27	RELATIONSHIP DETECTED
OTACDTCL 7 --	OTNUMDAY: 2.27	RELATIONSHIP DETECTED
01FTR24 --	01SEPLAN 23: 2.26	RELATIONSHIP DETECTED
01FTR24 --	OTACDTCL 13: 2.25	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 3: 2.24	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 9: 2.24	RELATIONSHIP DETECTED
01FTR11 --	01SEPLAN 23: 2.24	RELATIONSHIP DETECTED
OTACDTCL 13 --	OTAMTYPE 0: 2.23	RELATIONSHIP DETECTED
01SEPLAN 4 --	OTACDTLM: 2.22	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 12: 2.22	RELATIONSHIP DETECTED
01SEPLAN 23 --	OTACDTCL 4: 2.21	RELATIONSHIP DETECTED
01SEPLAN 18 --	OTACDTCL 5: 2.21	RELATIONSHIP DETECTED
01SEPLAN 14 --	01SEPLAN 19: 2.20	RELATIONSHIP DETECTED
01SEPLAN 23 --	OTAMTYPE 2: 2.19	RELATIONSHIP DETECTED
01AMTODU --	01FTR43: 2.19	RELATIONSHIP DETECTED
01FTR11 --	01SEPLAN 2: 2.16	RELATIONSHIP DETECTED
01SEPLAN 13 --	01SEPLAN 23: 2.15	RELATIONSHIP DETECTED
01SEPLAN 4 --	01SEPLAN 23: 2.15	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 15: 2.15	RELATIONSHIP DETECTED
01SEPLAN 20 --	01SEPLAN 23: 2.14	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 2: 2.13	RELATIONSHIP DETECTED

01AMAR30 --	01SEPLAN 6:	2.13	RELATIONSHIP DETECTED
01SEPLAN 17 --	OTACDTLM:	2.12	RELATIONSHIP DETECTED
01SEPLAN 18 --	01SEPLAN 22:	2.11	RELATIONSHIP DETECTED
01RCFCHG --	OTACDTCL 4:	2.10	RELATIONSHIP DETECTED
01ANPHON --	OTNUMDAY:	2.10	RELATIONSHIP DETECTED
01AMAR30 --	OTACDTCL 2:	2.10	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 3:	2.10	RELATIONSHIP DETECTED
01FTR04 --	01SEPLAN 11:	2.09	RELATIONSHIP DETECTED
01SEPLAN 13 --	OTACDTCL 4:	2.08	RELATIONSHIP DETECTED
01INTRAS --	OTACDTCL 11:	2.08	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 15:	2.08	RELATIONSHIP DETECTED
01ROMAIR --	01SEPLAN 4:	2.07	RELATIONSHIP DETECTED
01INTRAS --	OTACDTCL 4:	2.07	RELATIONSHIP DETECTED
01ANPHON --	01FTR41:	2.07	RELATIONSHIP DETECTED
01SEPLAN 12 --	OTACDTCL 2:	2.06	RELATIONSHIP DETECTED
01SEPLAN 10 --	01SEPLAN 24:	2.06	RELATIONSHIP DETECTED
01MINPEK --	01SEPLAN 1:	2.06	RELATIONSHIP DETECTED
01INTRAS --	01LCLLND:	2.06	RELATIONSHIP DETECTED
01SEPLAN 10 --	OTACDTCL 11:	2.04	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 10:	2.04	RELATIONSHIP DETECTED
01INTRAS --	01SEPLAN 24:	2.04	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 0:	2.04	RELATIONSHIP DETECTED
01SEPLAN 14 --	OTACDTCL 4:	2.03	RELATIONSHIP DETECTED
01MINPEK --	01SEPLAN 19:	2.03	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 0:	2.03	RELATIONSHIP DETECTED
01FTR19 --	01RCFCHG:	2.03	RELATIONSHIP DETECTED
01SEPLAN 23 --	01SEPLAN 25:	2.02	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 12:	2.02	RELATIONSHIP DETECTED
01SEPLAN 6 --	01SEPLAN 10:	2.02	RELATIONSHIP DETECTED
01MINPEK --	01ROMSRG:	2.02	RELATIONSHIP DETECTED
01AMAR30 --	01ROMAIR:	2.02	RELATIONSHIP DETECTED
01ACSCCHG --	OTACDTCL 5:	2.01	RELATIONSHIP DETECTED
01FTR24 --	01SEPLAN 2:	2.01	RELATIONSHIP DETECTED
01SEPLAN 18 --	OTACDTCL 14:	2.00	RELATIONSHIP DETECTED
01ROMAIR --	OTNUMDAY:	2.00	RELATIONSHIP DETECTED
01SEPLAN 13 --	OTAMTYPE 2:	1.99	RELATIONSHIP DETECTED
01ROMTOL --	OTNUMDAY:	1.98	RELATIONSHIP DETECTED
01AMAR60 --	OTACDTCL 1:	1.98	RELATIONSHIP DETECTED
01FTR04 --	01FTR13:	1.98	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 7:	1.97	RELATIONSHIP DETECTED
01FTR41 --	OTACDTCL 1:	1.97	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 1:	1.97	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 1:	1.96	RELATIONSHIP DETECTED
01ACSCCHG --	OTACDTCL 6:	1.96	RELATIONSHIP DETECTED
01SEPLAN 18 --	01SEPLAN 21:	1.94	RELATIONSHIP DETECTED
01FTR99 --	OTAMTYPE 2:	1.94	RELATIONSHIP DETECTED
01SEPLAN 10 --	OTACDTCL 12:	1.93	RELATIONSHIP DETECTED
01SEPLAN 10 --	01SEPLAN 19:	1.93	RELATIONSHIP DETECTED
01SEPLAN 9 --	OTACDTLM:	1.93	RELATIONSHIP DETECTED
01ROMTOL --	01SEPLAN 5:	1.93	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 13:	1.92	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 4:	1.92	RELATIONSHIP DETECTED
01INTRST --	OTACDTLM:	1.91	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 7:	1.91	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 18:	1.91	RELATIONSHIP DETECTED
01FTR24 --	01RCFCHG:	1.91	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 20:	1.90	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 14:	1.90	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 4:	1.90	RELATIONSHIP DETECTED
01FTR11 --	01FTR43:	1.90	RELATIONSHIP DETECTED
01INTRAS --	OTACDTLM:	1.89	RELATIONSHIP DETECTED
01FTR43 --	OTACDTCL 8:	1.89	RELATIONSHIP DETECTED
01SEPLAN 14 --	01SEPLAN 17:	1.88	RELATIONSHIP DETECTED
01ACSCCHG --	OTACDTCL 12:	1.88	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 10:	1.88	RELATIONSHIP DETECTED
01ROMAIR --	OTACDTCL 11:	1.87	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 15:	1.87	RELATIONSHIP DETECTED
01AMAR30 --	OTACDTCL 5:	1.87	RELATIONSHIP DETECTED
01FTR41 --	01INTRAS:	1.86	RELATIONSHIP DETECTED
OTACDTCL 4 --	OTACDTCL 12:	1.84	RELATIONSHIP DETECTED
01FTR99 --	01INTRAS:	1.84	RELATIONSHIP DETECTED
01INTRST --	01SEPLAN 7:	1.83	RELATIONSHIP DETECTED
01ROMAIR --	01SEPLAN 24:	1.82	RELATIONSHIP DETECTED

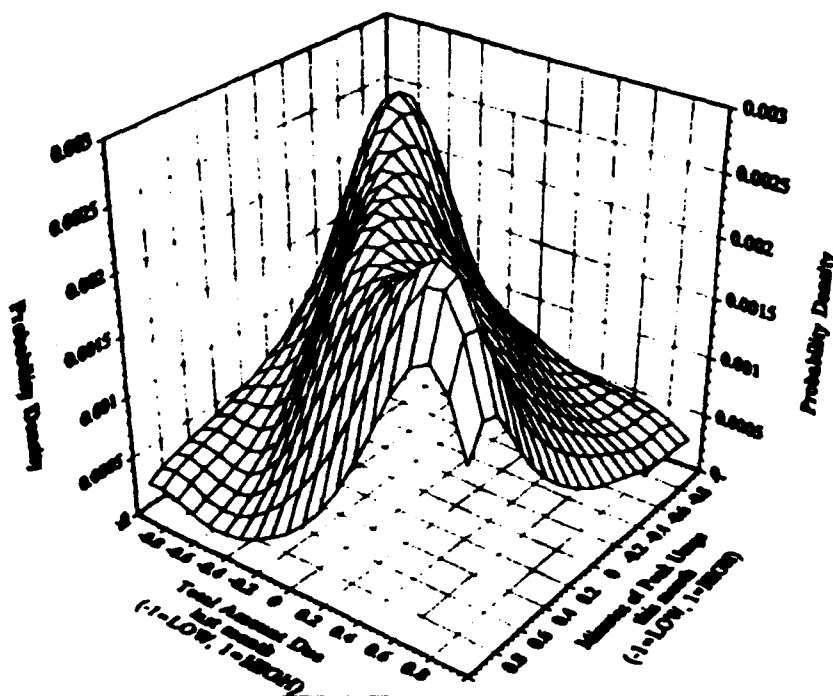
01INTRAS --	01SEPLAN 7:	1.82	RELATIONSHIP DETECTED
01ACSCHG --	OTACDTCL 10:	1.82	RELATIONSHIP DETECTED
01FTR10 --	OTACDTCL 8:	1.82	RELATIONSHIP DETECTED
01SEPLAN 5 --	OTAMTYPE 0:	1.80	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 25:	1.80	RELATIONSHIP DETECTED
01FTRT0T --	01SEPLAN 18:	1.80	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 3:	1.80	RELATIONSHIP DETECTED
01AMAR60 --	OTACDTLM:	1.80	RELATIONSHIP DETECTED
OTACDTCL 3 --	OTACDTCL 4:	1.79	RELATIONSHIP DETECTED
01FTR47 --	OTACDTCL 11:	1.79	RELATIONSHIP DETECTED
01SEPLAN 13 --	OTACDTCL 13:	1.78	RELATIONSHIP DETECTED
01FTR47 --	01SEPLAN 24:	1.78	RELATIONSHIP DETECTED
01FTR24 --	OTNUMDAY:	1.76	RELATIONSHIP DETECTED
01RCFCBG --	01ROMAIR:	1.75	RELATIONSHIP DETECTED
OTACDTCL 12 --	OTACDTCL 13:	1.74	RELATIONSHIP DETECTED
OTACDTCL 4 --	OTACDTCL 14:	1.73	RELATIONSHIP DETECTED
01INTRST --	01SEPLAN 13:	1.72	RELATIONSHIP DETECTED
01ACSCHG --	OTACDTCL 3:	1.71	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 11:	1.71	RELATIONSHIP DETECTED
OTACDTCL 3 --	OTACDTCL 13:	1.70	RELATIONSHIP DETECTED
01FTR13 --	OTNUMDAY:	1.69	RELATIONSHIP DETECTED
01FTR03 --	01FTR43:	1.68	RELATIONSHIP DETECTED
OTACDTCL 2 --	OTACDTCL 7:	1.67	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 7:	1.67	RELATIONSHIP DETECTED
01FTRT0T --	01SEPLAN 22:	1.67	RELATIONSHIP DETECTED
01SEPLAN 10 --	01SEPLAN 17:	1.66	RELATIONSHIP DETECTED
01FTRT0T --	OTACDTCL 14:	1.66	RELATIONSHIP DETECTED
01FTR43 --	01SEPLAN 0:	1.66	RELATIONSHIP DETECTED
01FTR41 --	OTACDTLM:	1.66	RELATIONSHIP DETECTED
01MINPEK --	OTACDTCL 4:	1.65	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 8:	1.65	RELATIONSHIP DETECTED
01FTR04 --	01SEPLAN 18:	1.65	RELATIONSHIP DETECTED
OTACDTCL 13 --	OTACDTCL 14:	1.64	RELATIONSHIP DETECTED
01INTRST --	OTNUMDAY:	1.64	RELATIONSHIP DETECTED
01FTRT0T --	01SEPLAN 19:	1.64	RELATIONSHIP DETECTED
01SEPLAN 25 --	OTACDTCL 8:	1.63	RELATIONSHIP DETECTED
01AMPHON --	01SEPLAN 11:	1.63	RELATIONSHIP DETECTED
01SEPLAN 0 --	OTACDTCL 4:	1.62	RELATIONSHIP DETECTED
01RCFCBG --	01SEPLAN 6:	1.62	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 3:	1.62	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 22:	1.61	RELATIONSHIP DETECTED
OTACDTCL 4 --	OTACDTCL 5:	1.61	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 24:	1.61	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 0:	1.61	RELATIONSHIP DETECTED
01FTR24 --	01FTR99:	1.61	RELATIONSHIP DETECTED
01SEPLAN 0 --	OTACDTCL 11:	1.60	RELATIONSHIP DETECTED
01SEPLAN 6 --	01SEPLAN 24:	1.59	RELATIONSHIP DETECTED
01FTR04 --	01LCLIND:	1.59	RELATIONSHIP DETECTED
01SEPLAN 6 --	OTACDTCL 11:	1.58	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 6:	1.58	RELATIONSHIP DETECTED
01SEPLAN 23 --	OTAMTYPE 0:	1.57	RELATIONSHIP DETECTED
01SEPLAN 5 --	OTAMTYPE 2:	1.57	RELATIONSHIP DETECTED
01SEPLAN 3 --	01SEPLAN 14:	1.57	RELATIONSHIP DETECTED
01AMAR60 --	01SEPLAN 17:	1.57	RELATIONSHIP DETECTED
OTACDTCL 7 --	OTACDTCL 8:	1.56	RELATIONSHIP DETECTED
OTACDTCL 4 --	OTACDTCL 10:	1.56	RELATIONSHIP DETECTED
01FTR24 --	01ROMAIR:	1.56	RELATIONSHIP DETECTED
OTACDTCL 10 --	OTAMTYPE 2:	1.55	RELATIONSHIP DETECTED
01SEPLAN 15 --	OTACDTCL 13:	1.55	RELATIONSHIP DETECTED
01AMAR30 --	01SEPLAN 23:	1.55	RELATIONSHIP DETECTED
01FTR99 --	OTACDTCL 3:	1.55	RELATIONSHIP DETECTED
01FTR41 --	01FTR47:	1.55	RELATIONSHIP DETECTED
01SEPLAN 4 --	OTACDTCL 8:	1.54	RELATIONSHIP DETECTED
01AMPHON --	01SEPLAN 18:	1.54	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 7:	1.54	RELATIONSHIP DETECTED
01SEPLAN 19 --	01SEPLAN 24:	1.53	RELATIONSHIP DETECTED
01SEPLAN 15 --	OTACDTCL 4:	1.53	RELATIONSHIP DETECTED
OTACDTCL 5 --	OTACDTCL 13:	1.53	RELATIONSHIP DETECTED
01SEPLAN 19 --	OTACDTCL 11:	1.52	RELATIONSHIP DETECTED
01SEPLAN 14 --	01SEPLAN 15:	1.52	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 19:	1.52	RELATIONSHIP DETECTED
01SEPLAN 6 --	01SEPLAN 19:	1.50	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 14:	1.50	RELATIONSHIP DETECTED

01ROMSRG --	01SEPLAN 5:	1.50	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 20:	1.50	RELATIONSHIP DETECTED
01FTR43 --	01SEPLAN 2:	1.50	RELATIONSHIP DETECTED
01FTR11 --	01SEPLAN 11:	1.50	RELATIONSHIP DETECTED
01SEPLAN 11 --	01SEPLAN 21:	1.49	RELATIONSHIP DETECTED
0TACDTCL 10 --	0TACDTCL 13:	1.48	RELATIONSHIP DETECTED
01SEPLAN 7 --	01SEPLAN 23:	1.47	RELATIONSHIP DETECTED
01SEPLAN 5 --	01SEPLAN 12:	1.47	RELATIONSHIP DETECTED
01FTR20 --	01MINPEK:	1.47	RELATIONSHIP DETECTED
0TACDTCL 8 --	0TAMTYPE 1:	1.45	RELATIONSHIP DETECTED
01RCFCHG --	01SEPLAN 3:	1.45	RELATIONSHIP DETECTED
01FTRTOT --	0TACDTCL 13:	1.45	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 22:	1.45	RELATIONSHIP DETECTED
01MINPEK --	01SEPLAN 15:	1.44	RELATIONSHIP DETECTED
01INTRST --	0TACDTCL 11:	1.44	RELATIONSHIP DETECTED
01AMAR30 --	01INTRST:	1.44	RELATIONSHIP DETECTED
01FTR99 --	0TAMTYPE 0:	1.44	RELATIONSHIP DETECTED
01SEPLAN 5 --	0TACDTCL 7:	1.43	RELATIONSHIP DETECTED
01AMAR30 --	01AMAR30:	1.43	RELATIONSHIP DETECTED
01SEPLAN 16 --	01SEPLAN 18:	1.42	RELATIONSHIP DETECTED
01SEPLAN 8 --	01SEPLAN 18:	1.42	RELATIONSHIP DETECTED
01ROMAIR --	0TAMTYPE 2:	1.42	RELATIONSHIP DETECTED
01INTRAS --	01SEPLAN 11:	1.42	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 13:	1.42	RELATIONSHIP DETECTED
01AMCUDU --	01SEPLAN 20:	1.42	RELATIONSHIP DETECTED
01FTR11 --	0TACDTCL 13:	1.42	RELATIONSHIP DETECTED
01INTRST --	01SEPLAN 24:	1.41	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 17:	1.41	RELATIONSHIP DETECTED
01SEPLAN 3 --	01SEPLAN 10:	1.39	RELATIONSHIP DETECTED
01ROMSRG --	01SEPLAN 4:	1.39	RELATIONSHIP DETECTED
01ROMAIR --	01SEPLAN 3:	1.38	RELATIONSHIP DETECTED
01MINPEK --	01SEPLAN 12:	1.38	RELATIONSHIP DETECTED
01INTRAS --	01SEPLAN 15:	1.38	RELATIONSHIP DETECTED
01ACSCHG --	01FTR85:	1.38	RELATIONSHIP DETECTED
01FTR43 --	01SEPLAN 18:	1.38	RELATIONSHIP DETECTED
01SEPLAN 13 --	01SEPLAN 14:	1.37	RELATIONSHIP DETECTED
01SEPLAN 12 --	0TNUMDAY:	1.37	RELATIONSHIP DETECTED
01SEPLAN 4 --	01SEPLAN 14:	1.37	RELATIONSHIP DETECTED
01RCFCHG --	01SEPLAN 4:	1.37	RELATIONSHIP DETECTED
01LCLLND --	01RCFCHG:	1.37	RELATIONSHIP DETECTED
01SEPLAN 14 --	01SEPLAN 20:	1.36	RELATIONSHIP DETECTED
01ROMTOL --	0TACDTCL 11:	1.36	RELATIONSHIP DETECTED
01LCLLND --	01SEPLAN 24:	1.36	RELATIONSHIP DETECTED
01FTR11 --	01INTRST:	1.36	RELATIONSHIP DETECTED
01SEPLAN 18 --	0TAMTYPE 2:	1.35	RELATIONSHIP DETECTED
01SEPLAN 10 --	01SEPLAN 15:	1.35	RELATIONSHIP DETECTED
01SEPLAN 3 --	0TACDTCL 8:	1.35	RELATIONSHIP DETECTED
01LCLLND --	0TACDTCL 11:	1.35	RELATIONSHIP DETECTED
01INTRST --	01SEPLAN 0:	1.35	RELATIONSHIP DETECTED
01FTR11 --	01SEPLAN 18:	1.35	RELATIONSHIP DETECTED
01SEPLAN 17 --	0TACDTCL 8:	1.34	RELATIONSHIP DETECTED
01SEPLAN 13 --	0TACDTCL 10:	1.34	RELATIONSHIP DETECTED
01SEPLAN 17 --	01SEPLAN 24:	1.33	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 10:	1.33	RELATIONSHIP DETECTED
01FTR85 --	0TACDTCL 1:	1.33	RELATIONSHIP DETECTED
01SEPLAN 2 --	01SEPLAN 7:	1.32	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 17:	1.32	RELATIONSHIP DETECTED
01ROMTOL --	01SEPLAN 24:	1.32	RELATIONSHIP DETECTED
01SEPLAN 17 --	0TACDTCL 11:	1.31	RELATIONSHIP DETECTED
01SEPLAN 11 --	0TACDTCL 10:	1.31	RELATIONSHIP DETECTED
01SESTRC 0 --	01SESTRC 3:	1.31	RELATIONSHIP DETECTED
01SEPLAN 12 --	01SEPLAN 23:	1.30	RELATIONSHIP DETECTED
01SEPLAN 6 --	01SEPLAN 17:	1.30	RELATIONSHIP DETECTED
01SEPLAN 14 --	01SEPLAN 25:	1.29	RELATIONSHIP DETECTED
01FTR47 --	01RCFCHG:	1.29	RELATIONSHIP DETECTED
01SEPLAN 3 --	0TACDTCL 2:	1.28	RELATIONSHIP DETECTED
01ROMTOL --	0TAMTYPE 2:	1.28	RELATIONSHIP DETECTED
01ROMTOL --	01SEPLAN 3:	1.28	RELATIONSHIP DETECTED
01FTR43 --	01INTRAS:	1.28	RELATIONSHIP DETECTED
01FTR13 --	0TACDTCL 13:	1.28	RELATIONSHIP DETECTED
01SEPLAN 6 --	0TACDTCL 10:	1.27	RELATIONSHIP DETECTED
01RCFCHG --	01SEPLAN 23:	1.27	RELATIONSHIP DETECTED
01RCFCHG --	01SEPLAN 7:	1.27	RELATIONSHIP DETECTED

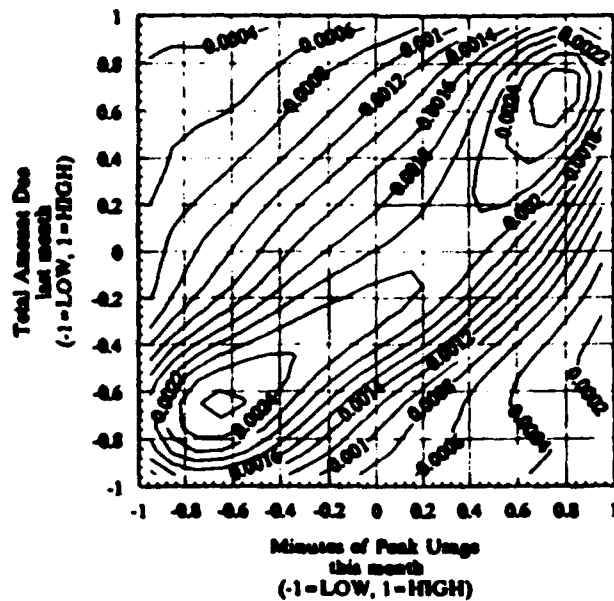
01ANPHON	--	01SEPLAN 19:	1.27	RELATIONSHIP	DETECTED
01FTR24	--	01SEPLAN 10:	1.27	RELATIONSHIP	DETECTED
01FTR24	--	01ROMTOL:	1.27	RELATIONSHIP	DETECTED
01FTR13	--	01SEPLAN 23:	1.27	RELATIONSHIP	DETECTED
01SEPLAN 1	--	OTACDTCL 3:	1.26	RELATIONSHIP	DETECTED
01AMSTAT 1	--	01ANPHON:	1.26	RELATIONSHIP	DETECTED
01AMSTAT 0	--	01ANPHON:	1.26	RELATIONSHIP	DETECTED
01ANPHON	--	01RCFCHG:	1.26	RELATIONSHIP	DETECTED
01FTR99	--	01MINPER:	1.26	RELATIONSHIP	DETECTED
01FTR24	--	01SEPLAN 14:	1.26	RELATIONSHIP	DETECTED
01SEPLAN 17	--	01SEPLAN 19:	1.25	RELATIONSHIP	DETECTED
01SEPLAN 12	--	OTAMTYPE 2:	1.25	RELATIONSHIP	DETECTED
01AMAR60	--	01SEPLAN 6:	1.25	RELATIONSHIP	DETECTED
01AMAR60	--	01FTR99:	1.24	RELATIONSHIP	DETECTED
01AMSTAT 1	--	01SEPLAN 11:	1.23	RELATIONSHIP	DETECTED
01AMSTAT 0	--	01SEPLAN 11:	1.23	RELATIONSHIP	DETECTED
01ROMATR	--	01SEPLAN 19:	1.23	RELATIONSHIP	DETECTED
01ANPHON	--	OTACDTCL 2:	1.23	RELATIONSHIP	DETECTED
01ANTODU	--	OTACDTCL 12:	1.23	RELATIONSHIP	DETECTED
01AMAR30	--	01SEPLAN 7:	1.23	RELATIONSHIP	DETECTED
01FTR41	--	OTACDTCL 14:	1.23	RELATIONSHIP	DETECTED
01FTR12	--	01SEPLAN 23:	1.23	RELATIONSHIP	DETECTED
01SEPLAN 10	--	01SEPLAN 13:	1.22	RELATIONSHIP	DETECTED
01SEPLAN 4	--	01SEPLAN 10:	1.22	RELATIONSHIP	DETECTED
01SEPLAN 10	--	01SEPLAN 20:	1.21	RELATIONSHIP	DETECTED
01SEPLAN 4	--	OTACDTCL 13:	1.21	RELATIONSHIP	DETECTED
01FTR03	--	OTACDTCL 4:	1.21	RELATIONSHIP	DETECTED
01SEPLAN 25	--	OTACDTCL 13:	1.20	RELATIONSHIP	DETECTED
01SEPLAN 16	--	OTACDTCL 1:	1.20	RELATIONSHIP	DETECTED
01SEPLAN 5	--	01SEPLAN 22:	1.20	RELATIONSHIP	DETECTED
01FTR0T	--	01ROMTOL:	1.20	RELATIONSHIP	DETECTED
01ANPHON	--	OTACDTCL 7:	1.20	RELATIONSHIP	DETECTED
01AMAR60	--	01SEPLAN 1:	1.20	RELATIONSHIP	DETECTED
01AMAR30	--	01SEPLAN 13:	1.20	RELATIONSHIP	DETECTED
01SEPLAN 2	--	OTACDTCL 3:	1.19	RELATIONSHIP	DETECTED
01AMSTAT 1	--	01FTR41:	1.19	RELATIONSHIP	DETECTED
01AMSTAT 0	--	01FTR41:	1.19	RELATIONSHIP	DETECTED
01AMAR60	--	01SEPLAN 14:	1.19	RELATIONSHIP	DETECTED
01AMAR60	--	01SEPLAN 11:	1.19	RELATIONSHIP	DETECTED
01AMAR60	--	OTAMTYPE 3:	1.18	RELATIONSHIP	DETECTED
OTACDTCL 4	--	01SEPLAN 12:	1.18	RELATIONSHIP	DETECTED
01SEPLAN 2	--	01SEPLAN 22:	1.18	RELATIONSHIP	DETECTED
01AMAR30	--	OTACDTCL 1:	1.18	RELATIONSHIP	DETECTED
01AMAR30	--	OTAMTYPE 0:	1.17	RELATIONSHIP	DETECTED
OTACDTCL 10	--	01SEPLAN 6:	1.17	RELATIONSHIP	DETECTED
01LCLND	--	01SEPLAN 25:	1.15	RELATIONSHIP	DETECTED
01SEPLAN 10	--	01SEPLAN 18:	1.15	RELATIONSHIP	DETECTED
01INTRST	--	01SEPLAN 13:	1.15	RELATIONSHIP	DETECTED
01FTR85	--	OTACDTCL 3:	1.14	RELATIONSHIP	DETECTED
01SEPLAN 23	--	01SEPLAN 1:	1.14	RELATIONSHIP	DETECTED
01AMAR30	--	OTACDTCL 8:	1.13	RELATIONSHIP	DETECTED
01SEPLAN 18	--	OTAMTYPE 2:	1.13	RELATIONSHIP	DETECTED
01SEPLAN 10	--	OTACDTCL 1:	1.13	RELATIONSHIP	DETECTED
01SEPLAN 8	--	01SEPLAN 24:	1.13	RELATIONSHIP	DETECTED
01SEPLAN 3	--	01SEPLAN 6:	1.13	RELATIONSHIP	DETECTED
01ROMSRG	--	01SEPLAN 10:	1.13	RELATIONSHIP	DETECTED
01AMAR30	--	OTACDTLM:	1.12	RELATIONSHIP	DETECTED
01SEPLAN 21	--	01SEPLAN 16:	1.12	RELATIONSHIP	DETECTED
01SEPLAN 11	--	01SEPLAN 11:	1.12	RELATIONSHIP	DETECTED
01SEPLAN 8	--	01SEPLAN 21:	1.12	RELATIONSHIP	DETECTED
01SEPLAN 5	--	OTACDTCL 11:	1.12	RELATIONSHIP	DETECTED
01SEPLAN 3	--	01SEPLAN 3:	1.12	RELATIONSHIP	DETECTED
01SEPLAN 0	--	01SEPLAN 13:	1.12	RELATIONSHIP	DETECTED
01INTRAS	--	OTAMTYPE 2:	1.12	RELATIONSHIP	DETECTED
01AMAR60	--	01FTR04:	1.12	RELATIONSHIP	DETECTED
01AMAR30	--	01SEPLAN 22:	1.12	RELATIONSHIP	DETECTED
01FTR41	--	01SEPLAN 6:	1.11	RELATIONSHIP	DETECTED
01SEPLAN 3	--	OTACDTCL 13:	1.11	RELATIONSHIP	DETECTED
01SEPLAN 1	--	OTACDTLM:	1.11	RELATIONSHIP	DETECTED
01AMSTAT 1	--	OTACDTLM:	1.11	RELATIONSHIP	DETECTED
01AMSTAT 0	--	01SEPLAN 1:	1.11	RELATIONSHIP	DETECTED
01ROMSRG	--	01SEPLAN 3:	1.11	RELATIONSHIP	DETECTED
01FTR99	--	01SEPLAN 9:	1.11	RELATIONSHIP	DETECTED
01FTR11	--				

01FTR10 --	01MINPER:	1.11	RELATIONSHIP DETECTED
01SEPLAN 18 --	OTAMTYPE 0:	1.10	RELATIONSHIP DETECTED
01SEPLAN 15 --	01SEPLAN 24:	1.10	RELATIONSHIP DETECTED
01SEPLAN 9 --	OTACDTCL 4:	1.10	RELATIONSHIP DETECTED
01SEPLAN 5 --	OTACDTCL 4:	1.10	RELATIONSHIP DETECTED
01ROMTOL --	01SEPLAN 14:	1.10	RELATIONSHIP DETECTED
01RCFCHG --	01SEPLAN 1:	1.10	RELATIONSHIP DETECTED
01INTRST --	01SEPLAN 5:	1.10	RELATIONSHIP DETECTED
01AMTODU --	01SEPLAN 20:	1.10	RELATIONSHIP DETECTED
01SEPLAN 21 --	OTACDTCL 12:	1.09	RELATIONSHIP DETECTED
01SEPLAN 21 --	OTACDTCL 8:	1.09	RELATIONSHIP DETECTED
01SEPLAN 15 --	OTACDTCL 11:	1.09	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 24:	1.09	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 15:	1.09	RELATIONSHIP DETECTED
01INTRAS --	OTNUMDAY:	1.09	RELATIONSHIP DETECTED
01SEPLAN 24 --	OTAMTYPE 2:	1.08	RELATIONSHIP DETECTED
01SEPLAN 11 --	OTACDTCL 8:	1.08	RELATIONSHIP DETECTED
01SEPLAN 6 --	01SEPLAN 15:	1.08	RELATIONSHIP DETECTED
01SEPLAN 2 --	OTACDTCL 12:	1.08	RELATIONSHIP DETECTED
01SEPLAN 1 --	OTACDTCL 11:	1.08	RELATIONSHIP DETECTED
01SEPLAN 0 --	01SEPLAN 1:	1.08	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 20:	1.08	RELATIONSHIP DETECTED
01FTR11 --	OTACDTCL 2:	1.08	RELATIONSHIP DETECTED
OTACDTCL 11 --	OTAMTYPE 2:	1.07	RELATIONSHIP DETECTED
01SEPLAN 22 --	01SEPLAN 23:	1.07	RELATIONSHIP DETECTED
01SEPLAN 19 --	OTACDTCL 4:	1.07	RELATIONSHIP DETECTED
01SEPLAN 3 --	01SEPLAN 19:	1.07	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 6:	1.07	RELATIONSHIP DETECTED
01INTRAS --	01SEPLAN 5:	1.07	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 4:	1.07	RELATIONSHIP DETECTED
01RCFCHG --	01SEPLAN 17:	1.06	RELATIONSHIP DETECTED
01RCFCHG --	OTAMTYPE 2:	1.06	RELATIONSHIP DETECTED
01ANPHON --	01SEPLAN 3:	1.06	RELATIONSHIP DETECTED
01AMAR60 --	OTACDTCL 4:	1.06	RELATIONSHIP DETECTED
01FTR99 --	01SEPLAN 21:	1.06	RELATIONSHIP DETECTED
01SEPLAN 0 --	OTACDTCL 7:	1.05	RELATIONSHIP DETECTED
01MINPER --	OTACDTCL 2:	1.05	RELATIONSHIP DETECTED
01MINPER --	01SEPLAN 0:	1.05	RELATIONSHIP DETECTED
01FTRTOT --	01SEPLAN 4:	1.05	RELATIONSHIP DETECTED
01AMAR60 --	01FTR47:	1.05	RELATIONSHIP DETECTED
01AMAR60 --	01FTR20:	1.05	RELATIONSHIP DETECTED
01SEPLAN 15 --	01SEPLAN 19:	1.04	RELATIONSHIP DETECTED
01SEPLAN 10 --	OTAMTYPE 0:	1.04	RELATIONSHIP DETECTED
01ACSCHG --	OTACDTCL 7:	1.04	RELATIONSHIP DETECTED
01ACSCHG --	01FTR43:	1.04	RELATIONSHIP DETECTED
01FTR99 --	OTACDTCL 5:	1.04	RELATIONSHIP DETECTED
01FTR41 --	01SEPLAN 21:	1.04	RELATIONSHIP DETECTED
01FTR24 --	01SEPLAN 0:	1.04	RELATIONSHIP DETECTED
01SEPLAN 1 --	01SEPLAN 19:	1.03	RELATIONSHIP DETECTED
01AMTODU --	OTACDTCL 3:	1.03	RELATIONSHIP DETECTED
01AMCUDU --	OTACDTCL 3:	1.03	RELATIONSHIP DETECTED
01AMCUDU --	01FTR10:	1.03	RELATIONSHIP DETECTED
01FTR11 --	01INTRAS:	1.03	RELATIONSHIP DETECTED
01ROMAIR --	01SEPLAN 7:	1.02	RELATIONSHIP DETECTED
01FTR24 --	01SEPLAN 6:	1.02	RELATIONSHIP DETECTED
01ROMSRG --	OTAMTYPE 2:	1.01	RELATIONSHIP DETECTED
01AMAR60 --	OTACDTCL 10:	1.01	RELATIONSHIP DETECTED
01AMAR30 --	01SEPLAN 17:	1.01	RELATIONSHIP DETECTED
01FTR42 --	01INTRST:	1.01	RELATIONSHIP DETECTED
01FTR10 --	01ROMTOL:	1.01	RELATIONSHIP DETECTED
01SEPLAN 21 --	01SEPLAN 23:	1.00	RELATIONSHIP DETECTED
01SEPLAN 18 --	OTACDTCL 10:	1.00	RELATIONSHIP DETECTED
01SEPLAN 13 --	01SEPLAN 24:	1.00	RELATIONSHIP DETECTED
01SEPLAN 4 --	01SEPLAN 24:	1.00	RELATIONSHIP DETECTED
01FTR04 --	01ROMTOL:	1.00	RELATIONSHIP DETECTED

Surface Plot of Probability Density Function



Contour Plot of Probability Density Function



x --	linear :	99.99
quadratic --	cosine :	53.04
x --	quadratic :	49.91
linear --	quadratic :	41.63
x --	cosine :	34.69
linear --	cosine :	31.18
cosine --	random :	12.79
linear --	random :	12.33
quadratic --	random :	12.18
x --	random :	7.78

Probing of K-space Relationship Discovery

Test Data Set 1

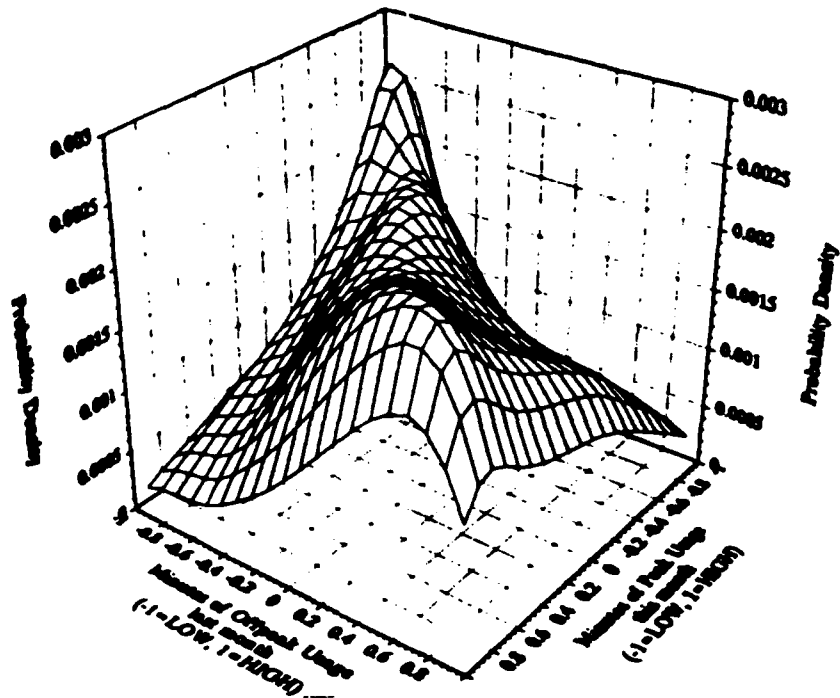
SN = 5

x --	linear :	99.99
quadratic --	cosine :	69.18
x --	quadratic :	53.38
linear --	quadratic :	51.26
linear --	cosine :	45.20
x --	cosine :	44.52
cosine --	random :	11.39
quadratic --	random :	10.27
linear --	random :	7.37
x --	random :	5.44

Probing of K-space Relationship Discovery

Test Data Set 1

SN = 10



Contour Plot of Probability Density Function

A contour plot showing the probability density function of Peak Usage (x-axis) and Offpeak Usage (y-axis). Both axes range from -1 to 1, with -1 representing LOW and 1 representing HIGH. The plot features several contour lines representing probability density values. The highest density (0.0022) is located in the upper right quadrant, indicating a high probability of high peak usage and high offpeak usage. Other contours show varying probabilities across the range of usage levels.

